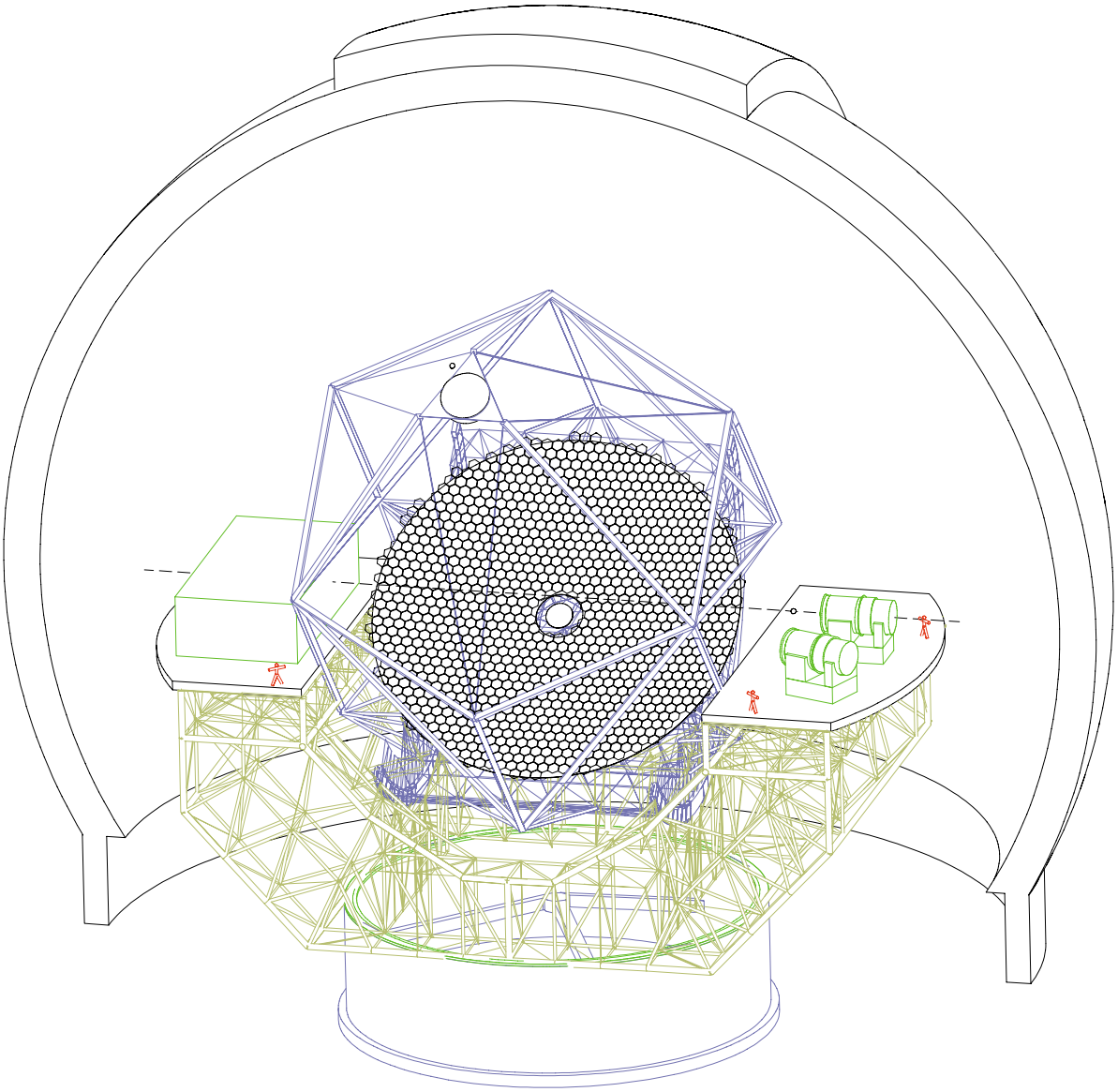
CELT Report Number 34

California Extremely Large Telescope

Conceptual Design for a Thirty-Meter Telescope

June 2002



University of California

California Institute of Technology

The California Extremely Large Telescope
Conceptual Design for a Thirty-Meter Telescope

For thousands of years astronomy has captured the imagination of all civilizations. In 1669 Vermeer vividly portrayed the astronomer's technology, discipline, and sense of awe.

Participants

Greg Ames (Blueline Eng.)
Jean Aubrun (ACSD)
Mike Bolte (UCSC)
Brian Bauman (LLNL)
Gary Brack (JPL)
Matthew Britton (CIT)
Jerry Cabak (UCO/Lick)
Gary Chanan (UCI)
Judy Cohen (CIT)
Al Conrad (WMKO)
Dave Cowley (UCSC)
Deborah Culmer (UCSC)
Richard Dekany (CIT)
George Djorgovski (CIT)
Richard Ellis (CIT)
Don Gavel (LLNL)
Andrea Ghez (UCLA)
Kirk Gilmore (UCO/Lick)
James Graham (UCB)
Raja Guhathakurta (UCSC)
Steve Gunnels (Paragon Eng.)
Stan Hermann (Vertex/RSI)
Bill Irace (JPL)
Richard Jared (LBNL)
Mike Jura (UCLA)
Shri Kulkarni (CIT)
Lee Laiterman (UCSC)
James Larkin (UCLA)
Hilton Lewis (WMKO)
Ken Lorell (ACSD)
Doug MacMartin (CIT)
Terry Mast (UCSC)
Keith Matthews (CIT)
Bruce Macintosh (LLNL)
Ian McLean (UCLA)
Maureen McLean (UCSC)
Steve Medwadowski (SJM Eng.)
Heather Mietz (UCO/Lick)
Joe Miller (UCSC)
Bob Minor (LBNL)
Richard Murray (CIT)
Jerry Nelson (UCSC)
Scot Olivier (LLNL)
Carol Osborne (UCSC)
Steve Padin (CIT)
Andreas Quirrenbach (UCSD)
Wal Sargent (CIT)
Alan Schier (Pilot Group)
Matthias Schoeck (UCI)
Gerry Smith (WMKO)
Gary Sommargren (LLNL)
Chuck Steidel (CIT)
Keith Taylor (CIT)
Tom Tombrello (CIT)
Mitchell Troy (JPL)
David Tytler (UCSD)
Steve Vogt (UCSC)
Dave Woody (CIT)

Preface

For many centuries the exploration of the world by courageous discoverers profoundly influenced the development of civilization. The discovery of new lands and new societies brought pivotal changes to our world. Whether it was the discoveries of Polynesian adventurers crossing the Pacific Ocean or European explorers in the New World, the desire to explore and understand the unknown is a fundamental trait of the human race. In many respects the scientists of the 20th century were the modern equivalent of these explorers. Venturing into the microscopic world of the atom and the vast reaches of the universe, they made discoveries that had as profound an impact on the development of civilization as those of the early explorers of our planet.

Entering the 21st century, we stand on the threshold of dramatic astronomical explorations. The universe has been a source of interest, inspiration, and wonderment since the earliest times. Today, newspapers and magazines are filled with exciting new discoveries of astronomy because they appeal to something fundamental in human nature. People are eager to know about the universe in which we live, how it began and how it evolved to its present state; and to understand our place in it. We want to know:

- Why is the space of the universe filled with galaxies?
- Why are galaxies filled with stars?
- Why are stars surrounded by planets?
- Is the existence of life an extremely rare event or common in the universe?

We are beginning to answer some of these questions with existing telescopes on the ground and in space. The next generation of astronomical instruments will dramatically improve our ability to find answers to these questions that have intrigued human beings for thousands of years.

Scientists at the University of California and the California Institute of Technology propose to design, build, and operate a 30-meter telescope that will be an extraordinarily powerful tool for exploring the universe. It will see farther into space and farther back in time than any instruments currently in use. It will give us unprecedented access to exquisite details of physical processes on both small and large scales and over most of the age of the universe.

- We expect to see galaxies at their birth, when the first stars formed in the universe and started the processes which resulted in the world on which we live.
- We expect to further understand the evolution of galaxies from birth to the present.
- We expect to have detailed views of stars and solar systems in the process of formation.
- We expect to observe directly planets in orbits around other stars, planets that may be the abodes of life.

All this and much more will be possible because of the enormous light-gathering power and extremely high spatial resolution of a 30-meter telescope. The same scientists who conceived of and very successfully led the creation of the two largest telescopes in the world at the Keck Observatory are now prepared and eager to lead this exciting new venture. We are confident it will succeed, and the discoveries that flow from this magnificent new instrument will advance our understanding of the universe to a profound new level.

Executive Summary

Following great success in the creation of the Keck Observatory, scientists at the California Institute of Technology and the University of California have begun to explore the scientific and technical prospects for a much larger telescope. The Keck telescopes will remain the largest telescopes in the world for a number of years, with many decades of forefront research ahead after that. Though these telescopes have produced dramatic discoveries, it is already clear that even larger telescopes must be built if we are to address some of the most profound questions about our universe. The time required to build a larger telescope is approximately ten years, and the California community is presently well-positioned to begin its design and construction. The same scientists who conceived, led the design, and guided the construction of the Keck Observatory have been intensely engaged in a study of the prospects for an extremely large telescope. Building on our experience with the Keck Observatory, we have concluded that the large telescope is feasible and is within the bounds set by present-day technology. Our reference telescope has a diameter of 30 meters, the largest size we believe can be built with acceptable risk. The project is currently designated the California Extremely Large Telescope (CELT).

CELT will have nine times the collecting area of a Keck telescope. This tremendous gain in light gathering ability will allow imaging and spectroscopy of the faintest and most distant known objects in the universe and will provide a powerful complement to any future space-based telescopes. Because of the travel time of light, a large telescope is a time machine that allows travel into the distant past by observing objects at great distances. The 30-meter telescope will allow scientists to study in detail for the first time the era when the matter of the universe first began to collect into the organized structures of stars and galaxies. Many of the characteristics of the world we see around us today, including the conditions for the existence of life, result directly from the processes that took place in the early universe. The 30-meter telescope will provide the crucial measurements needed for understanding this era. CELT will also allow the direct detections of planets orbiting other stars, observations that are crucial for our understanding of the formation of planetary systems that may well be the locations of extraterrestrial life. Many other important and exciting scientific questions can be addressed only by using a telescope of this size. **The 30-meter telescope will unquestionably have a profound impact on our understanding of the universe and our place in it.**

A major design goal of the project will be to make cost-saving improvements beyond those incorporated in the Keck telescopes. Based on our experience with Keck, a segmented primary mirror is unquestionably the best choice for a telescope of this size. We will study alternative segment polishing and segment support techniques with the aim of greatly reducing the cost of the optics. We will also investigate alternative structural designs, with the aim of reduced cost and improved optics-support performance. Adaptive optics aided by laser beacons will be used to remove the blurring effects of atmospheric turbulence. Adaptive optics will provide a remarkable increase in angular resolution, orders-of-magnitude increase in point source sensitivity, and the ability to detect and spatially resolve even the most distant galaxies. Some of the adaptive optics technology required for the 30-meter telescope is not currently available. However, its rapid development, aided in part by the new NSF Center for Adaptive Optics at UC Santa Cruz, is presently underway. Scientific instruments, such as cameras and spectrometers, are needed to analyze the light collected by the telescope. These instruments, while challenging to build, are feasible and central to the scientific exploitation of the telescope. In addition, we expect that close collaboration with industry will produce a high performance and cost-effective dome design. The site for the 30-meter telescope has not yet been selected. Candidate sites

will be researched in the near future, and will include Mauna Kea in Hawaii and several sites in Chile. If the telescope is located outside the United States, we expect that scientists in the host country will have some access to the facility.

During the past year we have carried out a conceptual design program for CELT. Several engineering design studies have created a solid basis for many of the assumptions in this proposal. We have studied and created conceptual designs for the telescope structure, telescope bearings and drives, stressing fixtures for segment polishing, interferometric testing of segment surfaces, and sensors and actuators for segment active control. Some of these studies have advanced well beyond the concept stage: We are now building prototype sensors, actuators, and a stressing fixture. We have detailed cost estimates for the fabrication and assembly of some key components.

Progress to date in the study of CELT has been extremely encouraging and clearly indicates that its successful completion is entirely feasible. There are no outstanding technical issues that would prevent this singular scientific opportunity from becoming a reality.

Chapter 1. Introduction and Overview

1.1 Introduction

Major advances in science have invariably followed major advances in the instruments that are used for scientific research. This is abundantly clear in the history of astronomy. Astronomy is a science driven by observations, and since the time of Galileo the growth of our knowledge of the nature and contents of the Universe has closely followed the development of larger and larger telescopes. The light-gathering power of optical telescopes increased steadily since Galileo's, up to that of the famous 5-meter Hale Telescope at Palomar Mountain (completed in 1948). However, for nearly fifty years following the completion of this telescope, no larger ones were constructed. The technology to make the large, high-quality mirrors reached its limiting size with the 5-meter. A new mirror technology was needed to take the next step. As a result of the creative ingenuity of scientists at the University of California and the California Institute of Technology, this hiatus in telescope technology advancement was finally ended with the construction of the twin Keck 10-meter telescopes. Using a dramatically different approach to large mirror fabrication, each 10-meter primary is made up of 36 smaller mirrors placed with very high precision into a mosaic that acts as a single large mirror. The great success of the Keck telescopes with their segmented primary mirrors demonstrates the outstanding capability of this new technology. It makes possible the next major step: the creation of a vastly more powerful telescope that will allow unprecedented views of the Universe.

1.2 The Critical Advantages of a 30-Meter Telescope: Light Gathering and Angular Resolution.

1.2.1 Light Gathering

Astronomy is an observational science, and the science that a telescope is capable of is directly related to the amount of light it collects. As the amount of light collected goes up, the amount of information that can be extracted from that light goes up as well. Much of the most important data obtained in astronomy is in the form of spectra; spreading the light into a spectrum considerably dilutes the intensity of light on the detector, making the amount of collected light even more important.

The promise of obtaining spectra from very faint objects is one of the principal motivations for building larger telescopes. For example, after the Hubble Space Telescope (HST) revealed the presence of many small, extremely faint objects in essentially any patch of sky, only the Keck telescopes could collect enough light so that spectra of these objects could be obtained. The spectra showed that many of these faint objects were exceedingly distant galaxies, so distant that the light collected had been traveling for as long as 10 billion years. Large telescopes act as time machines: As we look to great distances we are seeing into the distant past.

For the "Keck Planet Search," nearby stars are being surveyed for planetary systems. Although each star is bright enough to be observed with smaller telescopes, the tremendous light gathering power of the Kecks allows spectra to be obtained in which the light is highly dispersed in wavelength. This permits the measurement of extremely accurate velocities for the stars, and consequently, the tiny velocity variations that reveal the presence of planets. Although smaller telescopes have discovered some of the planets, using them is a very time consuming process. Progress in finding new planets has been tremendously accelerated using the Keck telescopes.

It is through the analysis of spectra that motions and chemical compositions of objects can be measured, the masses of galaxies can be determined, the presence of planets can be inferred, and a host of other

investigations can be made. Through its large increase in collecting area, a 30-meter telescope will enable the spectroscopic study of objects that are completely beyond the reach of any present telescope. Obtaining spectra for many of these objects, which can barely be detected on direct images with existing telescopes, is central to increasing our understanding of the Universe.

1.2.2 Angular Resolution and Adaptive Optics

At ground-based sites, turbulence in the earth's atmosphere blurs the light from astronomical sources. For the best sites this blurring is about 0.5 seconds of arc, regardless of telescope size. The Hubble Space Telescope, orbiting above the atmosphere, is able to make images with an angular resolution of 0.1 seconds of arc, a factor of 5 improvement. This increase in resolution alone has led to spectacular improvements in our understanding of a broad range of phenomena in the Universe. In the future, the highest angular resolutions will be achieved using very large ground-based telescopes and correcting for atmospheric blurring using the technology of adaptive optics.

For any telescope mirror there is a fundamental physical limit to the sharpness of the images it can produce. That limit is determined by the diameter (D) of the telescope and the wavelength (λ) of the light used. The angular image size is proportional to λ/D; so in principle, larger mirrors can produce smaller, sharper images. However, for ground-based telescopes the atmospheric turbulence blurs the images and prevents this limit from being reached. Fortunately, in recent years the new technology of adaptive optics has been developed. This powerful but complex technique provides a means of real-time correction of atmospheric blurring, allowing an angular resolution superior to that of HST to be obtained with the Keck 10-meter telescopes. Using adaptive optics, the limiting resolution of the Keck telescopes is four times better than that of HST, and a 30-meter telescope will improve on HST by a factor of 12.5. This increase in resolution, a greater improvement over HST than HST was over existing ground-based telescopes, is crucial for many forefront areas of investigation in astronomy and astrophysics.

The University of California at Santa Cruz was chosen in 1999 as the site for a $40M National Science Foundation Science and Technology Center for Adaptive Optics with Jerry Nelson (Project Scientist for CELT) as Director. There will be a tremendous synergy between the work of the Center and CELT.

1.3 CELT Complements Future Space-Based Telescopes.

Within the next 15 years NASA will likely launch the successor to the HST. The design currently under consideration has a primary mirror with a diameter of 6.0 meters. Because the atmosphere of the Earth absorbs light at certain wavelengths and also emits radiation strongly at other (infrared) wavelengths, there will always be strong motivation to have powerful telescopes above the atmosphere in space. However, even when a new telescope is in space, the motivation for larger ground-based telescope remains strong.

Due to the prohibitive costs of building, launching, and maintaining very large telescopes in space, ground-based telescopes will have significantly greater light-collecting capability than any space telescope, at least for the foreseeable future. This capability will enable the critical spectroscopic observations of very faint sources, a spectroscopy impossible with space telescopes. In addition, by using adaptive optics, the larger ground-based telescopes will yield significantly higher spatial resolution than will those in space.

1.4 The Observatory as a Center for Education.

A very important goal for the CELT Observatory will be to serve as a center for education at many levels. The CELT California headquarters and possibly a CELT Astrophysics Institute will host many educational functions for graduate and undergraduate students in astronomy, high school and grade school students with interests in science and technology, and the general public from all over the world. These functions will take the form of an extensive Web site, exhibits, lectures, workshops, and astronomy research resources. These activities will take place at observatory facilities, with the full support of Caltech and the University of California, whose dominant mission is education.

1.5 Overview of Document

This document in its entirety is both a description of the concepts developed in the CELT Conceptual Design Phase and a proposal for the Preliminary and Final Design Phases. It also includes our initial thoughts about the management and budget of the project as a whole.

Chapter 2 describes the scientific motivation for CELT. The projects described reflect the great breadth of science, ranging from solar system studies to investigation of the highest redshift universe, that CELT will be able to address.

Chapters 3 and 4 describe the telescope performance specifications and a conceptual optical design that will meet them. We have selected a three-mirror Ritchey-Chrétien optical design to provide multiple f/15 foci at two Nasmyth platforms. To save cost and reduce complexity we will not provide prime or Cassegrain foci.

Chapters 5 and 6 describe the primary, secondary, and tertiary mirrors. The design, fabrication, and assembly of the primary mosaic will be based on the successful experience with the Keck Observatory primary mirrors. Some modifications of the technology developed there will be made to reduce fabrication and maintenance costs.

Chapter 7 describes the telescope structure that supports the optics and points the optical axis of the system. For high structural stiffness to resist gravity and wind loads the mirror cell will be supported from below on two journals and the secondary mirror will be supported by a two-layered tensioned truss.

Chapter 8 describes the alignment of the other optical elements within the structure. Following the Keck experience we will build an alignment camera to provide the measurements to align the segments and optics, as well as to measure the surface shape of individual segments.

Chapter 9 gives the goals and conceptual designs for adaptive optics systems that will allow the optics to reach their intrinsic (diffraction-limited) angular resolution. Two distinct adaptive optics instruments are envisioned to conduct the majority of the science programs. These instruments are the low-order adaptive optics (LOAO) to perform optimized low-background science at mid-IR wavelengths (3-30 μm), and multi-conjugate adaptive optics (MCAO), which is optimized for near-infrared wavelength (1-2.5 μm) science. Specialized modes of AO operation, which may require dedicated instrumentation, can expand the available science beyond these two baseline AO capabilities. All these systems are discussed in Chapter 9.

Chapter 10 describes representative instrumentation that will use both seeing-limited and adaptive-optics-corrected images to address science goals. The proposed suite of instruments includes: low- and high-resolution optical spectrographs, a medium-resolution fiber-fed spectrograph, an AO camera, an AO deployable integral field spectrograph, and an extreme AO coronograph. These instruments and their scientific capabilities are described.

Chapter 11 gives initial telescope optical error budgets for the image quality with adaptive optics off and on. These will be used to determine the engineering tolerances for fabrication and assembly. They provide the quantitative basis for making cost, schedule, and performance trades in the design, fabrication, and assembly phases.

Chapter 12 describes the telescope enclosure and observatory support facilities.

Chapter 13 describes desired site characteristics, the method of site selection, and some candidate sites. We describe the measurements of weather and atmospheric properties required to make an informed site selection.

Chapter 14 gives some initial considerations for creating the extensive software and related hardware required for several aspects of the operating observatory. A well-defined global architecture and approach will be created from the beginning of the project.

Chapter 2. Scientific Motivation

2.1 Introduction

CELT will combine a nine-fold increase in collecting area over existing optical/IR telescopes with a five-fold improvement in spatial resolution (for the near-IR) over the proposed Next Generation Space Telescope (NGST). CELT will provide exciting "breakthrough" possibilities in most areas of observational astronomy from Solar System exploration (with some capabilities exceeding those of space probes), to beyond the edge of the currently-mapped universe. This chapter begins with discussions of relevant properties of sky brightness as a function of wavelength, and comparisons with current and proposed front-line facilities. The science opportunities for CELT are then discussed in detail for a subset of the areas listed below. Because of the 10-year time period between when these words are being written and the first-light of CELT, and because CELT will open unexplored parameter space, it is likely that the most exciting discoveries of CELT will be in areas we cannot anticipate at this time. Nevertheless, there are a number of forefront areas of astronomy that are currently sensitivity- or spatial-resolution-limited, for which CELT will be required for further progress. We list some of the areas where we anticipate CELT will play a major role. Those marked with an asterisk are discussed in more detail in this chapter. The science opportunities for CELT will be explored in considerably more detail in the next phase of the project, with detailed simulations allowing important feedback to the telescope and instrument designs.

Anticipated major CELT science areas:

- Solar System explorations*****
- extrasolar planet searches and studies (including direct detection of hot Jupiters)*****
- star formations*****
- stellar seismology: high precision structure determination for the full range of stellar spectral types
- high precision astrometry for faint sources; detailed in-situ 3-dimensional kinematics throughout the Galaxy and Local Group
- chemical evolution and star formation histories of galaxies to 30Mpc*****
- resolving the mysteries of the Galactic center
- active galactic nuclei and black hole demography*****
- supernovae beyond z = 1: cosmological tools and probes of the star formation history
- the distribution of mass in the universe: intergalactic medium and weak lensing studies to large redshift*
- gamma ray bursts throughout the visible universe
- the era of galaxy evolution: high spatial, moderate spectral resolution studies of faint galaxies at z = 1 - 5*****
- the end of the Dark Ages: near-IR investigations of the initial star and galaxy formation events at z = 5 - 15

Much of the science outlined in the following sections assumes that technology development and implementation techniques for adaptive optics will proceed rapidly during the next decade, with the CELT project among the leaders in advancing the field. Although the full scientific power of the observatory will not be deployed until adaptive optics systems are working at the level detailed in Chapter 9 of this document, we emphasize that seeing-limited (i.e., non-AO) observational capabilities in both the optical and near-IR will be advanced significantly by the 30-m aperture and state-of-the-art seeing-limited instruments (Chapter 10) anticipated for CELT. As a result, about half of the science projects described in this chapter could proceed uninhibited by a slip in either schedule or scope of providing AO systems by first light. We cite as an example that the Keck 10-m telescopes operated

without an AO system for the first seven years of normal operations, and most of its major scientific accomplishments to date have been based on seeing-limited observations. The advance in seeing-limited capability of CELT over Keck exceeds that of the advance of Keck over previous generation telescopes.

We believe that CELT will deliver AO systems performing to required specifications at first-light or soon thereafter; however, we emphasize that perceived and real technology risks in this area need not be seen as compromising the viability of the CELT observatory to significantly advance forefront science from the outset. Generally, the expectation is that adaptive optics capabilities will continue to improve during the lifetime of the CELT observatory and ultimately will exceed those assumed for the present science case.

2.2 Thirty-Meter Telescope Project Background

With increased light-gathering power and finer diffraction-limited images, optical/IR telescopes with larger mirrors will always benefit observational astronomy. Recent experiences with the Keck 10-m telescopes have shown how new facilities with larger apertures allow for "quantum leaps" in both the range and the quality of the resulting discoveries. The Keck light-gathering abilities allowed the discovery and study of populations of galaxies at z = 3 and beyond, made possible the discovery of the nature of gamma ray bursts, led to the discovery of the majority of the known extrasolar planets, was crucial for establishing the evidence of an accelerating universe, and revolutionized our understanding of the star formation history of the universe, to mention only a few of the dramatic Keck breakthroughs.

A pattern of breakthroughs can be discerned at other times throughout the 20th century, where each new generation of larger-primary-mirror telescopes has led to significant new astronomical discoveries. We cite two examples. In the 1920's the 2.5-m telescope at Mt. Wilson allowed the first measurement of the expansion of the universe. Within five years of the Palomar 5-m first light (in the late 1950's) our understanding of stellar populations in the Galaxy and other Local Group galaxies was advanced enormously, quasars were discovered, and the extragalactic distance scale was established to within 50% of the presently accepted numbers. In the past, the rationale for making advances in telescope aperture was almost exclusively driven by the larger collecting area; because delivered image size was set primarily by atmospheric turbulence, the smaller diffraction limit of larger primaries was generally not used to any scientific advantage. Increased angular resolution, as opposed to light-gathering power, is the primary reason that the Hubble Space Telescope (HST) has had such a profound impact on our understanding of the universe, despite its modest 2.5-m aperture. The HST spatial resolution of ~0.1 arcsec offered a ~5-fold improvement over previous ground-based facilities. With CELT, taking advantage of the recent and continuing revolution in adaptive optics, we will be able to make substantial improvements in both light-gathering capability and high spatial resolution simultaneously. In broad terms, the increase in capabilities from the Keck 10-m to the CELT 30-m will be similar to the angular resolution gain from ground-based facilities to HST in the mid-1990's, coupled with an historically unprecedented gain in light-gathering power, a factor of nine increase from Keck to CELT (compared to the factor of four from the Palomar 5-m to the Keck 10-m telescopes).

While the Keck Observatory continues to produce exciting new results, and will continue to do so for years to come, a host of important problems in astrophysics are already clearly beyond Keck's capabilities, whether because of inadequate spatial resolution, sensitivity, or both. Because the development of world-class astronomical facilities is a long lead-time activity, now is the time to think about the next step.

We envision CELT as an observatory with very broad capabilities operating over the wavelength range 0.3-30 μm, poised to address the most compelling science of its era. CELT will operate both with and without adaptive optics (AO). Unlike many past facilities, CELT is being designed with diffraction-limited performance as a driving force during a time in which full adaptive optics (AO) capabilities are being delivered on current 8-m-class telescopes.

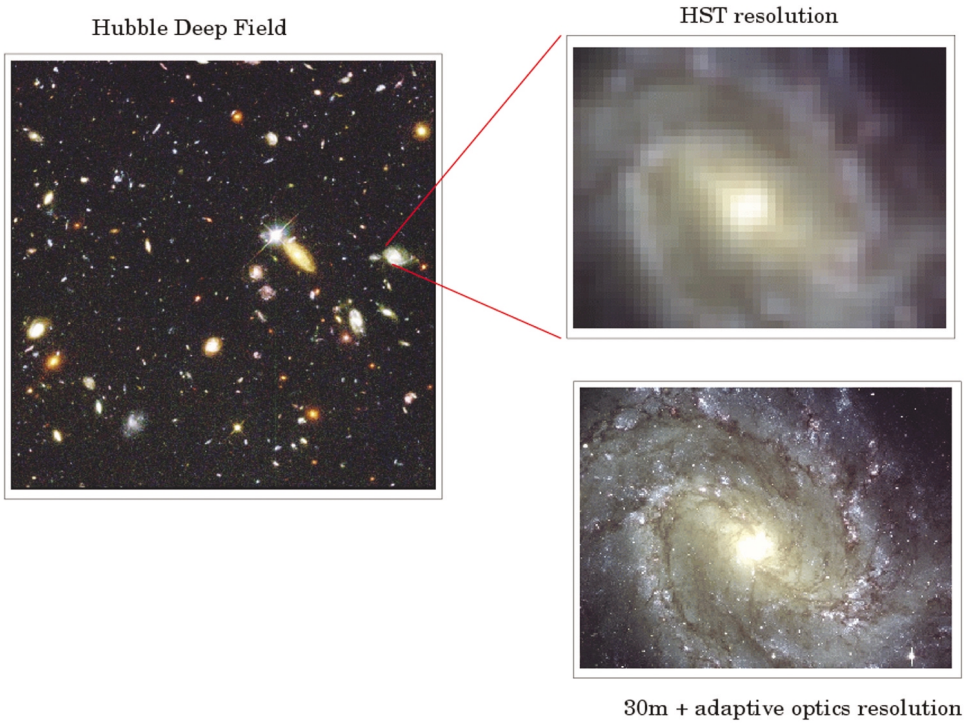


Figure 2-1. A simulation of imaging at the diffraction-limited CELT resolution versus the Hubble Space Telescope.

2.3 Technical Background for Science Motivation

In the sections describing the scientific investigations to be undertaken with CELT we will refer to various capabilities, often in comparison to existing or planned facilities. We have collected the information adopted for signal-to-noise ratio (S/N) calculations in this section. Although the CELT site has not yet been chosen, we will use the Mauna Kea sky as our fiducial for the background material on sky brightness and transmission. The principal variable between the sites under consideration is the typical amount of water vapor above the site.

2.3.1 Image Quality

A very significant advance in astronomical instrumentation is the realization of adaptive optics (AO) for near-real-time correction of wavefront distortions as light passes through the Earth's atmosphere. As has been demonstrated with the Keck 10-m telescopes, it is possible to achieve diffraction-limited images for wavelengths longer than ~ 0.9 μm over fields 20-40 arcsec in diameter. This advance brings the arena of high spatial resolution imaging back to ground-based facilities, as it will likely be the case for the next century that the largest telescopes will be built on the ground. If adaptive optics can

someday be extended to visual wavelengths, the space advantage for optical astronomy will all but disappear.

In the discussions that follow we will work in the context of two observing modes for the 30-m telescope. **Seeing-limited** observations will assume 0.5 arcsec FWHM images for point sources. This is typical in the visible for a good site and conservative for the near IR where image quality can be improved by as much as a factor of two compared to 500 nm. **Diffraction-limited** observations will assume the diffraction limit for a 30-m primary mirror. Each mode of observation is discussed in more detail below.

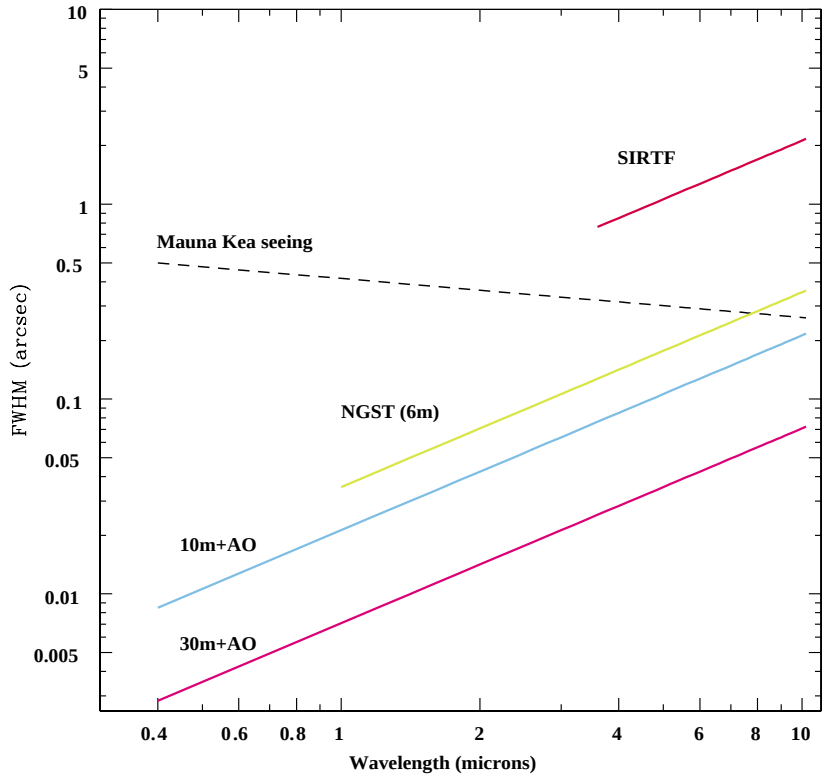


Figure 2-2. The diffraction limit as a function of wavelength for 10-m and 30-m telescopes, along with that of NGST and SIRTF. Seeing limited image size for Mauna Kea is shown as an example of image sizes at ground based telescopes. This curve applies equally to all large telescopes as long as the seeing size exceeds the diffraction limit of that telescope.

For reference in the following section, Figure 2-2 shows the FWHM of images for the Space Infrared Telescope Facility (SIRTF, a space-based 0.85-m infrared-optimized telescope), the currently-planned NGST (6 m, $l < 0.6$ μm), and diffraction-limited 10-m and 30-m telescopes. Also shown is typical ground-based seeing at Mauna Kea scaled by $l^{-0.2}$ assuming 0.5 arcsec images at 0.4 μm. Note that with currently-available AO systems it is not possible to correct the atmosphere for a 10-m (or 30-m) telescope to the diffraction limit with high Strehl for wavelengths shortward of 0.9 μm with large values of the Strehl ratio. It is our goal for CELT to extend useful correction (Strehl > 0.1) down to 0.5 microns (see Figure 3-1).

2.3.2 Atmospheric Transmission

The atmosphere is essentially opaque shortward of 0.3 μm and transparent up through the first significant water absorption band at 0.9 μm. In the near-IR, the commonly used bands (J, H, K) are set by the transmission of the atmosphere and are somewhat water vapor dependent. In the "thermal-IR," between 2.5 and 25 μm, the available ground-based observational bands become increasingly more dependent on water vapor content of the atmosphere. Approximately half of the wavelengths between 0.9 and 25 μm are essentially inaccessible from the ground. The following three figures (2-3, 2-4, and 2-5) show the near-IR and mid-IR atmospheric transmission for different values of opacity, usually expressed in terms of the effective column of precipitable water above the telescope site. A very good site for the thermal IR has a median opacity of $\simeq$ 1 mm (e.g., Mauna Kea, Chajnantor). The numbers are extremely well correlated with altitude, with the highest sites (> 4000 m) being much drier.

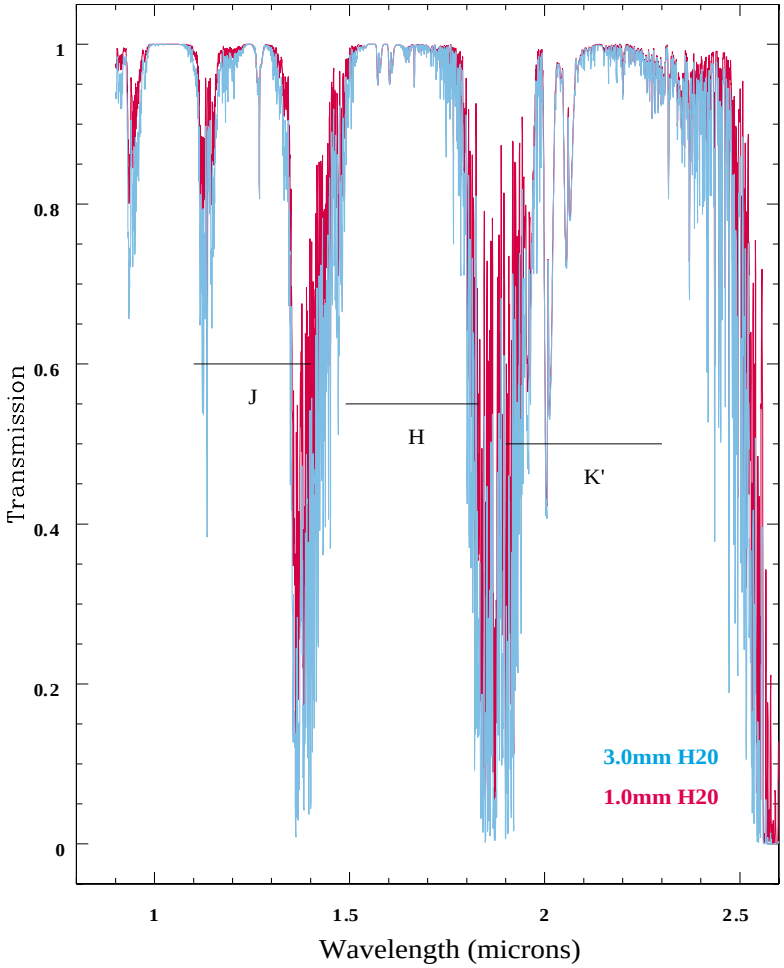


Figure 2-3. Atmospheric transmission in the near-IR for two water vapor levels.

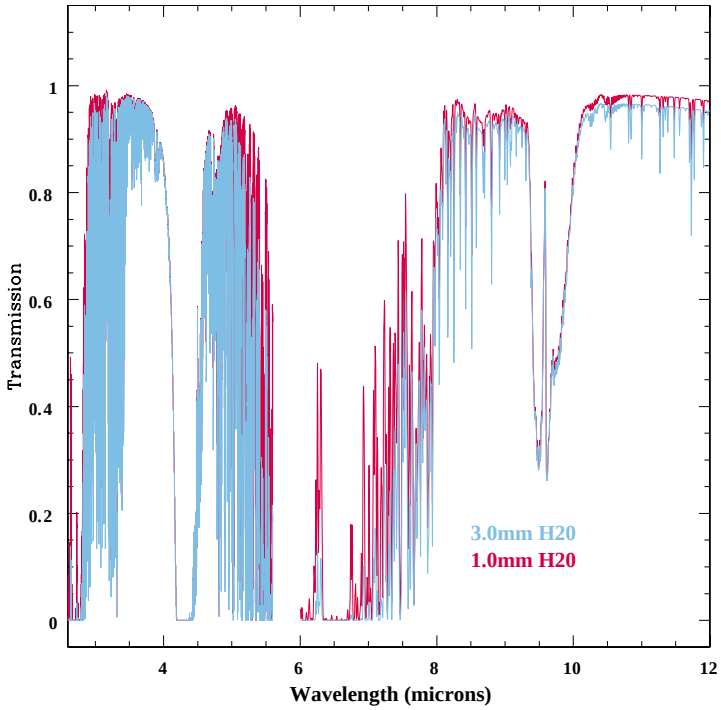


Figure 2-4. Atmospheric transmission in the thermal-IR for two water vapor levels.

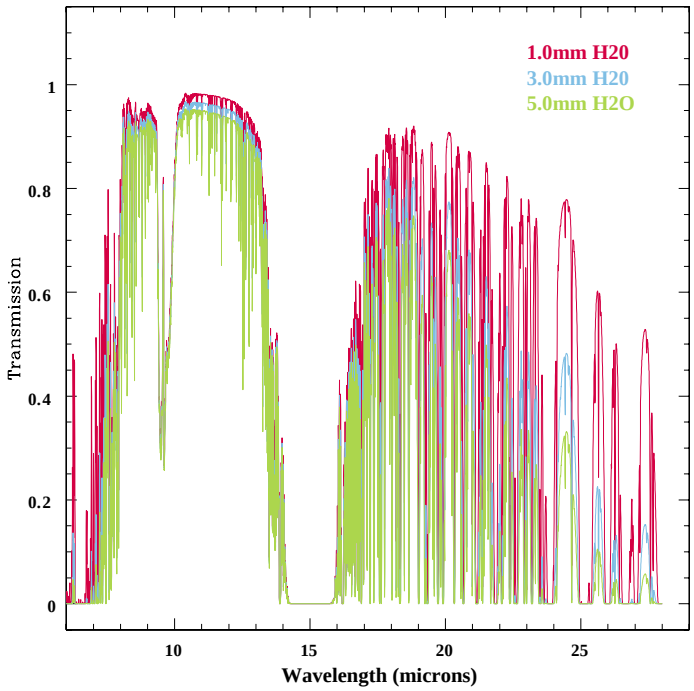


Figure 2-5. Atmospheric transmission in the 7-30 μm range for water vapor levels ranging from 1 mm to 5 mm; 5 mm is typical for sites at altitudes lower than 4000 m.

2.3.3 "Sky" Brightness

The table below gives the background sky brightness (from all sources not local to the telescope and enclosure) through commonly used broadband filters. These numbers are measured at the Canada France Hawaii Telescope on Mauna Kea and will be assumed for the rest of this chapter. Tabulated numbers are for new moon and the zenith. From ~ 0.7 μm through ~ 2 μm the night sky emission is dominated by OH molecular emission. At longer wavelengths water vapor and thermal continuum are the dominant source of photons. The telescope design becomes very important for background levels beyond 5 μm.

Table 2-1. Broadband sky brightness for Mauna Kea

Band	**Central λ**	**Brightness**		**Flux**	
	(μm)	(mag arcsec^{-2})	(AB mag arcsec^{-2})	(μJy arcsec^{-2})	(photon cm^{-2}s^{-1}μm^{-1}arcsec^{-2})
U	0.36	21.6	22.5	3.3	1.74 x 10^{-2}
B	0.44	22.3	22.2	4.8	1.76 x 10^{-2}
V	0.55	21.1	21.1	13.2	3.62 x 10^{-2}
R	0.64	20.3	20.6	20.9	5.50 x 10^{-2}
I	0.79	19.2	19.7	47.9	1.02 x 10^{-1}
J	1.23	14.8	15.6	2089.3	2.49
H	1.66	13.4	14.7	4786.3	4.20
K	2.22	13.5	15.4	2511.9	1.74

For spectral observations we will use the emission models shown in the following figures, each based on measurements at Mauna Kea.

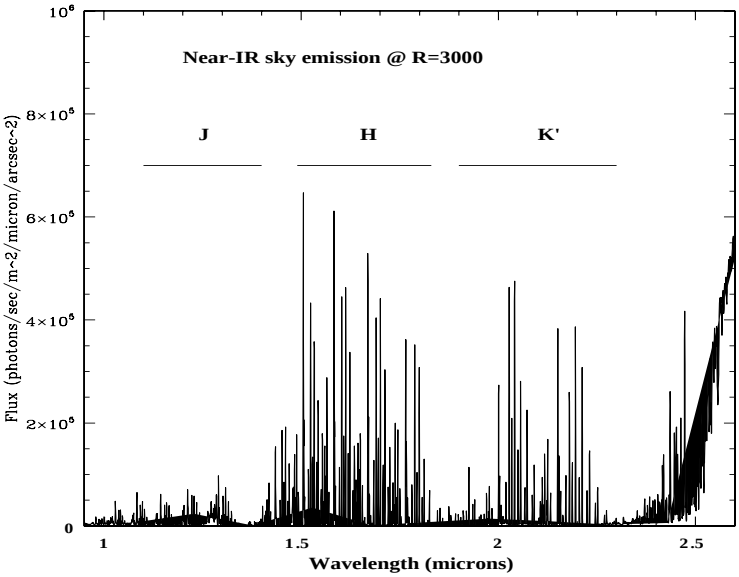


Figure 2-6. The emission from the atmosphere in the near-IR. Note that the level of the continuum emission is much lower than the average when all of the narrow OH emission features are included. Most spectroscopic observations will assume that the spectral resolution is high enough that only ~ 5% of the spectrum is contaminated by strong OH lines (R ≥ 4000). The broadband sky brightness in the J, H, and K bands are reduced by approximately a factor of 100 (5 magnitudes) between the bright OH lines.

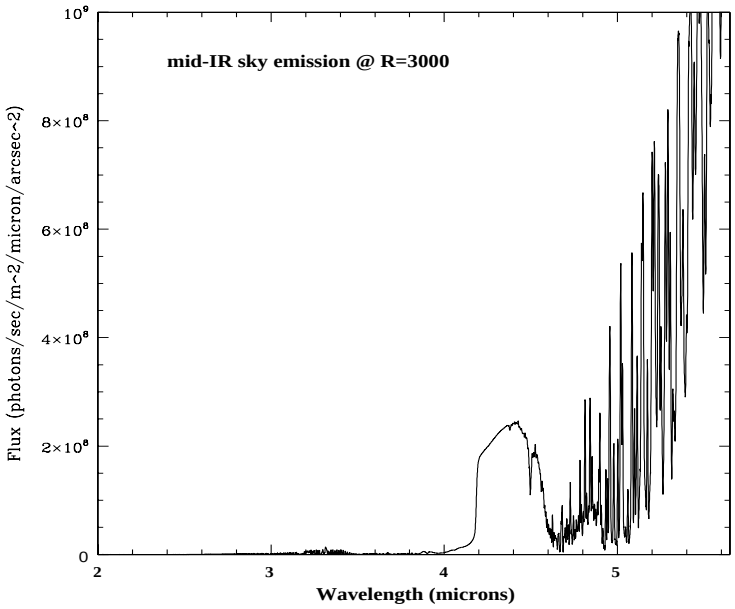


Figure 2-7. The emission from the atmosphere in the IR. The intensity axis is plotted linearly to emphasize the tremendous increase in atmospheric emission longward of 4 μm.

As we discuss below, achieving good spectroscopic sensitivity in the near-IR depends on having high enough spectral resolution that the effective sky background is that of the near-IR continuum, between the narrow OH features.

2.4 General Performance Capabilities of a 30-Meter Telescope

For the following discussion we will use this definition of signal-to-noise ratio (S/N).

$$S/N = S\,t / [\,S\,t + (sky \times At) + D\,A\,t + RN\,A\,]^{1/2} \qquad (2\text{-}1)$$

where "S" is the detected photon rate from the source, "Sky" is the detected photon rate per detector pixel from all foreground and background sources coincident on the sky with the source, "D" is the dark current rate per pixel of the detector; "RN" is the readout noise per pixel of the detector; "t" is the exposure time; and "A" is the number of pixels in the aperture in which the detection is measured. The expression above assumes that the estimate of the local sky does not contribute significantly to the variance (i.e., the background is estimated over a significantly larger area than the object of interest).

2.4.1 Thirty-Meter versus Ten-Meter Telescope

CELT will be dramatically faster and more capable than the best ground-based optical/IR telescopes. The comparison between a 30-m and 10-m with similar detectors and at the same site is very straightforward. Both S and Sky scale with telescope primary diameter D_M. For the case of AO correction to the diffraction limit, $A \sim D_M^{-2}$. In the common case where the sky background is the dominant noise source (faint sources and long exposure times), the limiting flux at fixed S/N and exposure time scales as D_M^{-1}, a factor of 3 comparing CELT to the Keck 10-m telescopes. The exposure time to reach a

given S/N scales like D_M^{-2} for the seeing-limited case, so that CELT is nine times faster than Keck. For diffraction-limited observations of unresolved sources, the gain of CELT over a Keck telescope is a remarkable factor of 81. In the diffraction-limited case for resolved objects, the gain will fall somewhere between a factor of 9 and 81 depending on the size of the object.

2.4.2 CELT and the Next Generation Space Telescope (NGST)

NGST is a NASA mission that is in the development stages, intended for launch in the 2009-2010 time frame, with a total cost, including technology development, of about $2 billion. The current design concept is a 6-m aperture, lightweight deployable primary mirror, designed to be diffraction-limited at 2 μm, optimized in the 1-5 μm range, but zodiacal-background (as opposed to telescope thermal background) limited to 10 μm. The combination of the envisioned detectors and gold–coated optics will limit the range in the visible to longward of 0.6 μm. At the time of this writing, the planned NGST instruments are a 1-5-μm imager, a 1-5-μm multi-object spectrograph, and a mid-IR combination imager and low-resolution spectrograph that will likely work to wavelengths of 25-30 μm.

CELT will outperform NGST at all wavelengths < 2.5 μm, and at longer wavelengths when the higher angular resolution of CELT is needed or helpful. In general, CELT and NGST are complementary because each will excel at a different type of observation.

In the comparison with space-based telescopes, the large decrease in sky brightness for telescopes above the atmosphere is a significant factor. At optical wavelengths, the background sky at V is only slightly brighter on the ground, but starting around 0.9 μm where OH emission from the atmosphere becomes important, the ground-based sky becomes orders of magnitude brighter (reaching a factor of 10^6 by 3.5 μm). Final design of the NGST telescope will set the background levels achieved longward of 2.5 μm (see Figure 2-8).

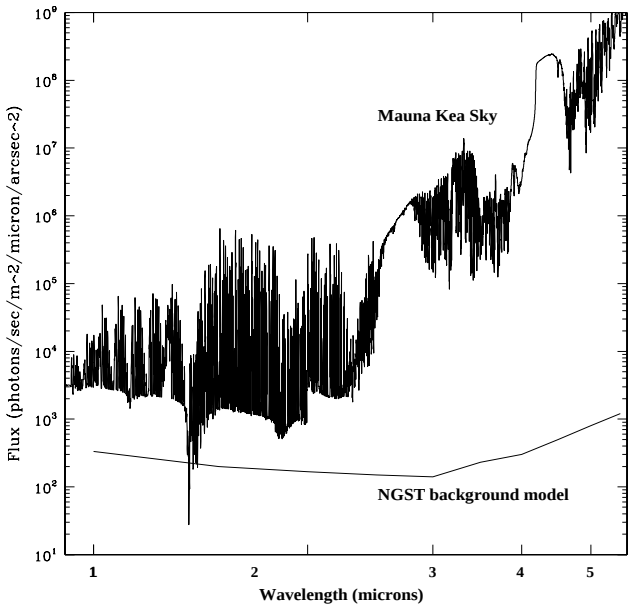


Figure 2-8. The dark sky comparison between Mauna Kea and a model for the NGST. There is a component of the spacecraft emission in this model longward of 3 μm.

Most of the relatively bright background in the 1-2.5 μm ("near-IR") windows for observations from the ground is concentrated into very bright emission from OH molecules in the upper atmosphere, with very narrow intrinsic line widths; as a result, the effective background at these wavelengths can be reduced to within a factor of ~10 of that achieved in space by observing at high spectral resolution. Instruments using "integral field," or 3-dimensional spectroscopic techniques (2-dimensional imaging with high-resolution spectroscopic information recorded simultaneously), effectively suppress the OH emission by censoring wavelength channels contaminated by atmospheric emission lines. *This same high spectral resolution is the main scientific requirement for unraveling the physical details of faint astrophysical sources, and it will require a very large collecting area (aperture) and very large instruments that are much better suited to terrestrial environments.*

Beyond 2.5 μm, NGST will be the telescope of choice for both imaging and low-to-moderate dispersion spectroscopy; the exception is when the ~5 times higher spatial resolution achieved with CELT would more than outweigh the significant loss in sensitivity inherent in terrestrial sites. As described below, there are significant areas of astronomy where thermal-IR observations with CELT will be revolutionary even in the era of NGST.

Figure 2-9 is adapted from Gillette and Mountain (1998) and shows the relative S/N for NGST versus CELT for three different spectral resolutions in the near- and mid-IR bands accessible from the ground. Red bars are for spectral resolution R =10,000, green bars for R =1000 and blue bars for R = 5 (broadband imaging). (Note: At present, NGST will not have spectroscopic capability with R > 1000. This is largely because of the gains to be had using large-aperture ground-based telescopes.) The combination of low background in space and relatively high dark current for IR detectors means that at the higher resolution, NGST observations would be detector-noise-limited. For this plot, NGST exposures of 4 x 1000 seconds and S/N of 10 were assumed. The detector characteristics for CELT and NGST assumed are: dark current = 0.01 e^-/s, read noise = 4 e^- for wavelengths shortward of 5.5 μm; and dark current = 10 e^-/s, read noise = 30 e^- for observations longward of 5.5 μm.

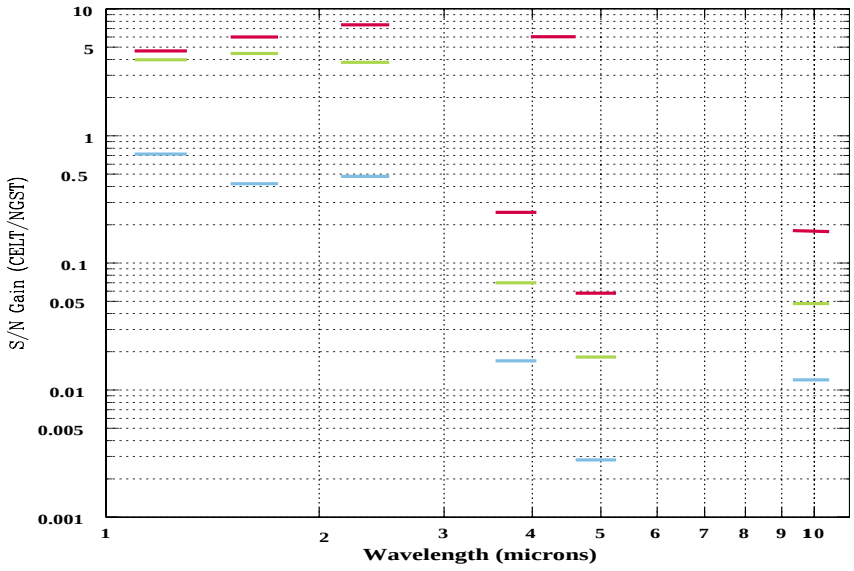


Figure 2-9. S/N gain for CELT versus a 6.5-m NGST for three spectral resolutions. Red bars are for R=10,000, green for R=1000, and blue for R=5.

At longer wavelengths, Figure 2-10 shows the expected limiting fluxes for R = 10 observations for CELT, Keck, SIRTF, and different configurations for NGST. Also shown is the background at two angles with respect to the ecliptic for the zodiacal light background level.

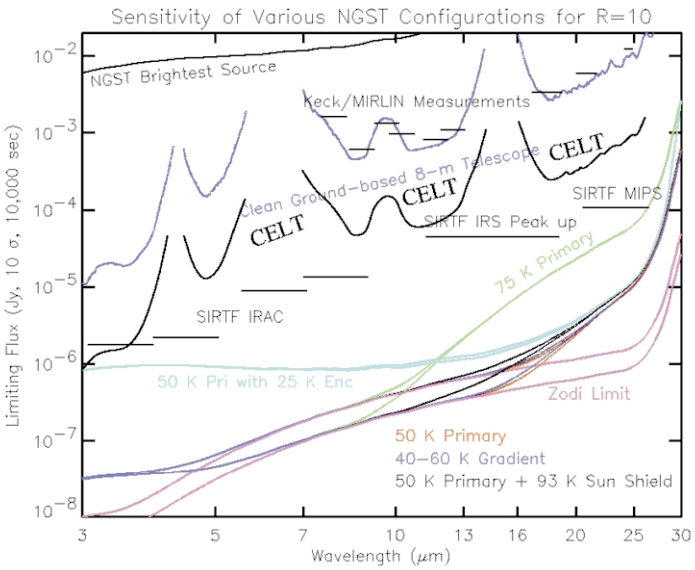


Figure 2-10. CELT, Keck, NGST and SIRTF at long wavelengths. Note that CELT will have close to the mid-IR sensitivity of SIRTF (but with 35 times higher spatial resolution) and will approach that of NGST at 25 μm if the specifications on telescope temperature are relaxed to 75K. These calculations were done assuming an 8-m NGST (current de-scoped telescope is 6 m) and a CELT emissivity of 5% (approximately the same as the Keck telescope).

Perhaps the greatest advantage of a ground-based facility over a space mission like NGST is the opportunity to carry out major programs over extended periods of time, and to continue to develop state-of-the-art instrumentation over a much longer observatory lifetime. This would allow a more rapid reaction to new developments in science over time, and would offer the versatility of applications both foreseen and unforeseen. Ground-based telescopes are much less expensive to build and operate and can be readily upgraded with better instrumentation as technology advances allow. One benefit of this situation is that ground-based facilities can be more flexible, and with a proper suite of instruments be far more versatile than an orbiting telescope. To a very large extent, NGST is being developed to excel in just the areas that will be difficult from the ground: extremely sensitive IR imaging and low-resolution identification-quality spectroscopy.

In short, the capabilities of NGST and a large (~30 m) ground-based telescope that is diffraction-limited would be almost completely complementary. This is discussed in more detail below on a scientific case basis.

2.4.3 CELT and ALMA

The Atacama Large Millimeter Array (ALMA) is a planned international facility (current partners are the U.S., Europe, and Japan in roughly equal proportion) that will consist of 64 12-m antennae placed on a high plateau (5000 m) in the Chilean Andes. On the current schedule, ALMA will be operational

by the end of this decade, and so would be a contemporary with CELT. ALMA will operate in atmospheric windows from 350 μm to 8 mm, optimized for wavelengths of ~1 mm. ALMA will operate in a number of different array configurations, ranging from a compact, nearly filled aperture array with baseline of 150 m, to a high-resolution configuration with maximum baseline of 10 km. In the compact configuration, ALMA will have point source sensitivity (estimated) for continuum observations of ~0.1 mJy (10 sigma, 1 hour) at 1 mm with a spatial resolution of ~0.5 arcsec; in the high resolution configuration ALMA will have a spatial resolution of 0.030 arcsec at 1 mm. As with NGST, CELT observations will complement those of ALMA. ALMA will excel at measuring thermal emission from dust, while CELT will observe the less obscured stars in the same local star-forming regions or high redshift galaxies.

The science case for ALMA includes many of the same fundamental questions we address below for CELT. For example, we argue that CELT will allow detailed physical investigations of galaxies at high redshift, using both the light-gathering power and the very high spatial resolution (roughly five times better at 1 μm than for ALMA at 1 mm). ALMA will be most powerful for examining the re-radiated emission from dust in high redshift galaxies: At 1 mm and z = 3, ALMA will be sensitive to thermal dust emission for galaxies exceeding $10^{11}L_{Sun}$ (star formation rates of $20M_{Sun}$ per year) , which would in the present-day universe be classified as luminous infrared galaxies (LIRGs); more luminous objects will probably be required for rest-frame far-IR spectroscopy. CELT, on the other hand, will have *spectroscopic* sensitivity in the UV to near-IR for objects down to the 10 nJy level (and down to rest-UV luminosities of perhaps $10^{8}L_{Sun}$, or unobscured star formation rates of ~$0.2M_{Sun}$ per year) at the same redshifts. It is now well known from observations in the sub-mm from the ground that there is a significant number of very luminous ($L_{bol} > 10^{12}L_{Sun}$), heavily obscured ($L_{FIR}/L_{UV} \sim 500$) sources at high redshift, but that the objects which comprise most of the sub-mm background are objects with about 10-times smaller luminosities and 10-times smaller extinction. These more common objects, which will be within the simultaneous reach of both CELT and ALMA, are those that produce most of the stars and metals in the high redshift universe. ALMA will allow the robust measurement of their bolometric energy production and molecular and atomic chemistry for the brightest sources, while CELT will more easily provide measures of redshifts, kinematics, and stellar populations.

ALMA and CELT will also be highly complementary for studies of the details of star formation in the nearby universe. With CELT operating in the 1-30 μm range at resolutions from 0.006 arcsec to 0.180 arcsec and ALMA in the 350 μm to 8 mm range with spatial resolution of 0.010 arcsec to 0.230 arcsec, these great observatories will resolve down to AU length scales at distances of 100-150 pc, providing access to the detailed chemistry and kinematics in the nearest star-forming regions in the galaxy. These measurements will be used together to understand the formation of stars and the protoplanetary disks that give rise to planets.

2.4.4 CELT and Other Future Facilities

We have learned from the Keck Observatory that an extremely important role is played by large ground-based optical/IR telescopes in following up sources first identified at other wavelengths, both on the ground and in space. For example, Keck has so far played a vital role in the identification of faint X-ray sources with Chandra, and gamma-ray bursts with Compton Observatory; worked with HST in identifying the high redshift supernovae, allowing measurement of the acceleration of the universe; spectroscopically identified rare low-mass stellar objects from the 2-MASS all-sky survey; identified most of the 850 μm sources for which redshifts are currently known; and obtained high-quality spectra

of the highest redshift QSOs from the Sloan Digital Sky Survey, spectra which is beginning to elucidate the epoch when the universe became fully-ionized. While much of the credit for these discoveries often goes to the other facilities, the science would be much less rich, and even not possible, without the Keck telescopes. We anticipate that CELT will play a similar role alongside future facilities and surveys both on the ground and in space. A large fraction of astronomy in the future will require spectroscopic observations of sources that are extremely faint in the optical and IR, and the sensitivity of large telescopes on the ground at these wavelengths will not be easy to surpass.

2.5 CELT Science Opportunities

There is virtually no astrophysical problem for which CELT will not represent a huge gain over the Keck telescopes, due to the order-of-magnitude gain in both collecting area and diffraction-limited PSF-size. Below we outline scientific programs that would particularly benefit from CELT; past experience shows that the science we envision now may not be among the most exciting projects for which CELT will be used by the time it is operational. Nevertheless, they illustrate of the kinds of gains that will be enabled using the next-generation state-of-the art ground-based optical/IR telescopes, and they point out the general capabilities desired for the telescope and instruments to guarantee a large scientific return in the future. Toward this end, we call out particular telescope and instrument goals/ requirements that would be necessary to carry out each of the proposed projects.

The science projects described are not meant to be exhaustive lists of all the areas where CELT will revolutionize observational astronomy. Rather, our intent is to reflect the great breadth of science, ranging from Solar System studies to investigation of the highest redshift universe, that CELT will be able to address.

2.5.1 Solar System Science with CELT

High resolution imaging

In many cases, global infrared images of planets and satellites of the Solar System observed with CELT would be *higher resolution* than those obtained by spacecraft exploring the Solar System. In addition, ground-based telescopes offer the possibility of significantly higher spectral resolution than has been obtainable on board spacecraft. The combination of these two capabilities will allow a 30-m telescope anchored to the Earth to make significant contributions to the exploration of the Solar System. A further advantage of CELT over explorer-type missions is the ability (thanks to the permanent nature of the facility and a routinely operating adaptive optics system) to monitor changes, e.g., weather and volcanic activity, on Solar System objects.

As an example, we consider the case of Jupiter's satellite Europa. Europa's surface is covered with water ice, but evidence suggests that underneath this ice layer a global liquid water ocean may exist. The water from this ocean may sporadically reach the surface of Europa in the many cracks penetrating the icy surface of the satellite. One piece of supporting evidence for this ocean is that low-resolution spectroscopy from the Galileo spacecraft has suggested that the dark regions around the cracks are composed of hydrated salts evaporated from the seawater below. If this were true, the composition of the salts would hold important answers to questions of composition of the proto-solar nebula, the degree of aqueous processing of the satellites, and the potential for supporting life or pre-organic chemistry. Unfortunately, at the spectral resolution of Galileo (R ~ 200), the identification of the dark materials on Europa is not certain. A resolution approximately 10 times higher, however, would allow the many different salt species or other possible components to be readily discerned. While such

spectral resolutions are routinely available from the ground today, at the low *spatial* resolution of typical ground-based observations the spectra of the large icy regions hide the spectra of the unresolved dark areas. At CELT resolution, however, the dark regions on Europa are resolved (Figure 2-11).

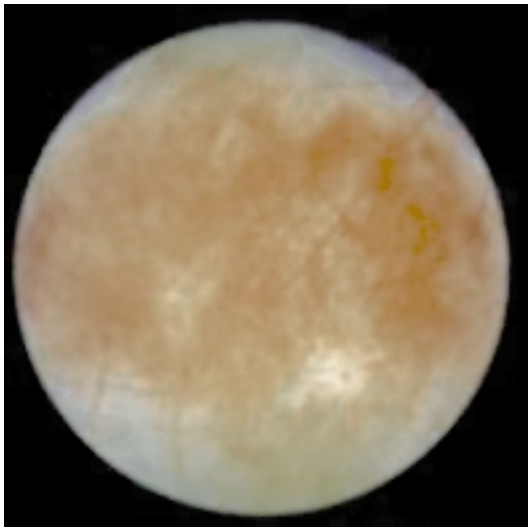


Figure 2-11. A visible-light Galileo image of Europa, convolved to the resolution of CELT. Linear cracks, expected to be the location of evaporated oceanic salts, are clearly resolved, as are craters and large icy regions. High-resolution spectroscopy of these features will allow definitive compositional understanding that is currently not possible.

High spatial and spectral resolution imaging of the satellite will allow definitive compositional identification that will help to solve many of the questions of this satellite and its possible oceanic interior. Similar problems will be solvable on the other Galilean satellites and on many other bodies of the Solar System.

Studying the edge of the Solar System

Most of the original material in the disk of gas, dust, and ice that formed the sun and planets of our Solar System has been heated, stirred, and compressed beyond recognition, leaving little information about the initial conditions that led to the current Solar System. Recently, however, planetary astronomers have discovered a vast swarm of small icy bodies -- named the Kuiper belt -- orbiting at the edge of the Solar System. While closer to the sun everything was heated and swept into planets, beyond Neptune the density of material was so low that no planets formed. These Kuiper belt objects (KBOs) have been preserved in deep freeze since the time of the formation of the Solar System. Study of the composition of these objects provides direct access to the make-up of the material out of which the planets formed.

The composition of icy bodies such as these is best determined through moderate-resolution (R ~ 1000) spectroscopy in the near-infrared (1-2.5 μm) where most important ices have strong absorption features. Because of their vast distances and small sizes, these objects are extremely faint (typically V ~ 24), so such infrared spectroscopy has not been possible. Using the Keck telescope, a few KBOs have been observed at lower resolution sufficient to detect ices with particularly strong and wide absorptions similar to water, but because of the small numbers of objects studied no concrete conclusions have been possible. In lieu of spectroscopy, astronomers have been studying the broadband colors of

KBOs from the blue to the infrared. While colors cannot provide compositional information, they can at least indicate which objects might be compositionally similar and which different. From studies of dozens of objects, it is apparent that KBOs come in a wide range of compositions with colors varying from essentially neutral to the reddest objects ever observed in the Solar System. It is clear that once spectroscopy is possible, astronomers will be rewarded with a rich assortment of spectral and compositional types holding many clues to the earliest history of the Solar System.

With the advances of laser guide star AO on the Keck telescope, the brightest KBOs will be just within reach of infrared spectroscopy. While these are likely to hold many compositional surprises, the color information suggests that these largest objects are the *least* compositionally different, and to understand the true diversity of compositions in the outer Solar System we will have to be able to reach to the much more abundant fainter objects. CELT will allow us to make this jump. With CELT we expect that hundreds (if not thousands, by then) of moderately faint KBOs will be well within the range of moderate-resolution spectroscopy. Because of the relative youth of this field, it is difficult to speculate on the discoveries that will be enabled by these advances. However, it is clear that this type of basic exploration of the Solar System will yield important insights into the formation of our and other planetary systems for many years to come.

Technical/Instrumental requirements
Most Solar System observations do not require particularly specialized instrumentation. Imaging of Solar System objects will be enabled with any of the planned AO imaging capabilities. Most important for efficient spectroscopy is the ability to perform small field (~ 2 arsec) integral field spectroscopy at moderate resolution (R ~ 1000) with large wavelength coverage to quickly cover the entire available band (cross dispersion is ideal).

The telescope needs to efficiently guide and track at non-sidereal rates as high as those expected for typical bright comets. Solar System observations are often time-specific, focusing on a certain face of a planet or alignment of satellites, so the telescope needs to maintain maximum flexibility in pointing and scheduling. Planets and satellites are often bright; care needs to be taken to make sure that no instrument is designed in such a way as to preclude observation of such bright objects.

2.5.2 Terrestrial Planet Searches and Studies with CELT

We describe potential observational programs with CELT to study terrestrial planetary systems outside of our Solar System. The large aperture of CELT will enable the order-of-magnitude leap required to advance from detecting and studying Jovian planets to similar investigations of terrestrial planets. Once extrasolar terrestrial planets are identified, we will be able to investigate spectroscopically with CELT whether life may exist on them.

Background
The first planets outside the Solar System were discovered around the pulsar PSR 1257+12 (Wolszczan and Frail 1992). The relatively short periods of 70 to 100 days for these planets implied that they were not long-lived survivors of the pre-supernova because they would have been engulfed when that star was a red supergiant. Instead, these pulsar-orbiting planets were viewed as a remarkable and curious consequence of a supernova explosion, but not in any obvious sense relevant to providing clues to the formation and evolution of our own Solar System.

The discovery in 1995 (Mayor and Queloz 1995) of a planet around the solar-type star 51 Peg, and the many other planetary discoveries since then, have dramatically improved our ability to learn about our own history. Currently, about 50 stars are known to possess roughly Jupiter-mass planets; at least 6% of approximately solar-mass main sequence stars have planetary companions (Marcy, et al., 2000). As we study these objects we are now in the position to learn more about the evolution of our own Solar System and even to address the question of whether life has formed and evolved elsewhere in the universe.

The bulk of our knowledge of planets comes from observing the reflex motion of the star that they orbit. We therefore learn about their orbital periods and eccentricities as well, M/sin(i), where M is the mass of the planet whose orbital plane is tipped at angle i relative to the plane of the sky. To date, the lowest known planetary masses are perhaps as low as 25% that of Jupiter (Marcy, Butler and Vogt 2000). The orbital periods of the planets are less than about 3 years.

In addition to observing their subtle gravitational effects on the star they orbit, there have been efforts to directly or indirectly detect light from the planets. The most striking result to date is the occultation of HD 209548 by its companion in a 3.5 day orbit (Henry, et al., 2000; Charbonneau, et al., 2000). The amplitude of the eclipse allows for a direct determination the planet's radius, which equals 1.5 ± 0.1 times the radius of Jupiter (Jha, et al., 2000).

If a planet is near enough to its host star, it can reflect enough light that its spectral lines might be detectable. By searching for such reflected light, Charbonneau, et al., (1999) have placed an upper limit of 0.3 to the albedo near 4800 Å of the planetary companion to τ Boo. This value of the albedo is lower than that found for the giant planets in the Solar System which are 0.46, 0.39, 0.60 and 0.58 for Jupiter, Saturn, Uranus and Neptune, respectively (Karkoschka 1994).

One of the strongest motivations for studying extrasolar planets is to learn if life exists elsewhere in the universe. While there is some possibility that Jovian planets, and especially their satellites, could harbor life (Sagan and Salepter 1976), it seems more promising to search for life on Earth-like planets.

In the Solar System, there are two classes of planets: the Jovian gas giants and the Earth-like (terrestrial) planets. The Jovian planets are 320 (Jupiter), 95 (Saturn), 15 (Uranus) and 17 (Neptune) times more massive than the Earth. Therefore, to study Earth-like systems, we should hope to study planets with masses not more than 10 times that of the Earth. This advance requires roughly an order-of-magnitude increase in sensitivity over current technology: this is achieved with the ratio of the collecting area of a 30-m CELT compared to a 10-m Keck.

During the past five years, we have succeeded in the ancient dream of detecting planets around other stars. We are beginning the physical study of these objects. Below, we list how CELT can play a vital role by expanding our studies to terrestrial as well as Jovian planets. First, we will describe how CELT can be used to detect new planets. Second, we will discuss how CELT can be used to study the planets that have been detected.

The Search for Terrestrial Planets

One of the main avenues of planetary research will be to identify more of these systems. With the ability to make catalogs, it will be possible to constrain models for the formation and evolution of these systems. Many questions arise:

- What is the mass distribution of planets? At the moment, it looks like the number of planets increases approximately as $dN/dM \sim M^{-1}$ (Marcy and Butler 2000). This implies that there are many low mass planets yet to be discovered.

- There is a hint that planets are mainly found around stars with metallicities greater than or equal to that of the Sun (Gonzalez, Wallerstein and Saar 1999). Is this an important clue for planet formation?

- Why do all the planets with a semi-major axis of their orbit larger than 0.2 AU have distinctly non-zero orbital eccentricities (Marcy and Butler 2000)?

- What is the chemical composition of these planets? Are the massive planets similar to Jupiter and Saturn?

- Three companions have been identified around HD 9826 (Marcy, Butler and Fischer 1999); are multiple planets common?

To answer these questions one must increase the sample of known planets.

Currently, the most successful technique to find extrasolar planets has been that of very high-precision radial velocity measurements; a typical rms scatter of 3 m/s has been achieved at Keck (Vogt, et al., 2000) and 7 m/s at European Southern Observatory (Santos, et al., 2000). To date, the surveys at Keck have focused on main sequence stars from spectral type F7 to M5 that have estimated masses of 1.2 M_{Sun} to 0.2 M_{Sun}, respectively (Drilling and Landolt 1999). The lowest mass star with a known planet is Gliese 876, which has a stellar mass of $0.32 \pm 0.3\ M_{Sun}$ while its companion has a mass of $2.1\ M_{Jup} / \sin(i)$ (Marcy, et al., 1998).

While current telescopes are very successful at finding planets around relatively bright stars, CELT will be able to search for planets around fainter stars. We expect CELT will have a high-resolution optical spectrograph similar to the HIRES echelle spectrograph at Keck. We therefore imagine that with an aperture of 30 m CELT will be able to study objects that are about 2.5 mag fainter than the current limit for Keck observations. Since Keck is currently observing stars as faint as $m_V = 11$ mag, this implies that CELT will be able to study stars as faint as about $m_V = 13.5$ mag. The advantage of extending the search to fainter stars is that it is possible to search for planets around lower mass stars. As a result, since the minimum mass to detect a planet scales directly to the mass of the orbited star, if we can monitor lower-mass stars, we can hope to find lower-mass planets.

Because the number distribution of planets seems to increase toward the lower masses, the current data suggest that there are large numbers of terrestrial planets in the Milky Way. With CELT, we should be able to identify terrestrial planets around nearby M dwarfs. Since planets with masses as low as 0.25 M_{Jup} have been detected around stars of 1 M_{Sun}, by extending the survey to stars of 0.15 M_{Sun}, it may be possible to detect planets with masses as low as 0.04 M_{Jup}, which is 13 M_{Earth}. With a large sample of surveyed stars, it should be possible with current technology to detect planets with masses perhaps a few times that of the Earth, if CELT is built (see Figure 2-12).

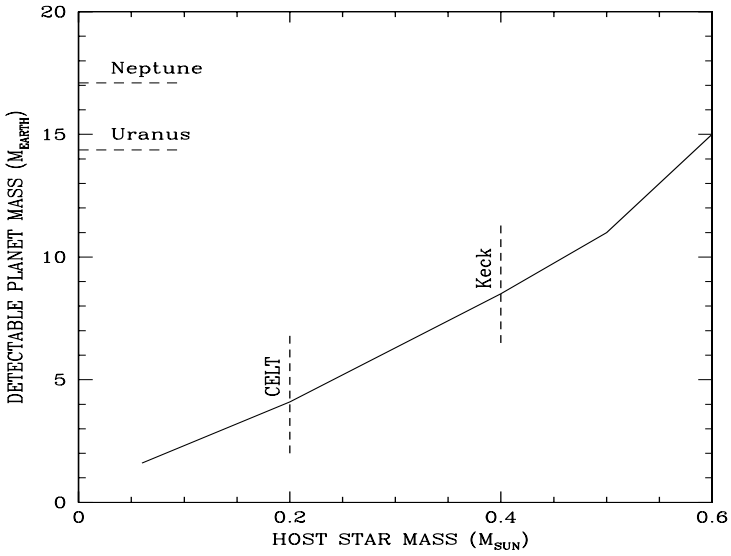


Figure 2-12. A plot of the minimum-mass planet that can be detected around a main-sequence host star, via radial velocity measurements, with the assumptions that the planet is a black body at 300 K, the orbital plane is viewed edge-on, and the star moves in a circular orbit with a speed of at least 3 m/s. (We choose T = 300 K so that the planet lies in the habitable zone and thus may possess life.) The dashed line for Keck is given by the constraint that the host star must be brighter than $m_V = 11.0$ mag. In the solar neighborhood, only stars earlier than M3 or about 0.4 M_{Sun} are found which satisfy this criterion. The dashed line for CELT assumes the same sensitivity, except now stars as faint as $m_V = 13.5$ mag can be studied. The larger telescope will enable searches for planets around stars as late as M5 or about 0.2 M_{Sun}. The Keck telescope appears to be restricted to searching for planets in the habitable zone that are $\geq$ 9 M_{Earth}, somewhat less than the mass of Uranus (14 M_{Earth}). With CELT, it may be possible to detect terrestrial planets of 3 M_{Earth} that lie in the habitable zone.

Assuming a random distribution of stars within the nearest 25 pc around the Sun, the number of sources that can be detected to a limiting flux, F, varies as $F^{-1.5}$. Because CELT can be used to study stars that are ~ 10 times fainter than is possible with Keck, then the sample of very low mass stars that can be studied with CELT is about 30 times larger than the sample that can be studied with Keck. For example, the catalog of nearby stars is maintained on the NSTARS web site (http://web05.arc.nasa.gov/nstars/). Among the 100 stars within 7.2 pc of the Sun, for stars brighter than $m_V = 11.0$ mag, there are two stars (Gliese 699 and Gliese 729) with estimated masses less than 0.2 M_{Sun}. However, for $m_V < 13.5$ mag, there are 30 stars with such low estimated masses.

Spectroscopic Study of Terrestrial Planets

If we are successful in detecting terrestrial-mass planets around nearby M dwarfs, then it may be possible to investigate spectroscopically the atmospheres of these planets, although the light from the central star will generally exceed that of the planet by a substantial amount.

As a representative example, consider a main-sequence star of mass 0.15 M_{Sun} with a luminosity, L_*, of 3×10^{-3} M_{Sun} yr^{-1}, an effective temperature, T_*, of 3200 K and a radius, R_*, of 1.3×10^{10} cm (Burrows, et al., 1993). Assume a planet of mass 3 M_{Earth} that lies at a distance, D, of 1.0×10^{11} cm from the star. Assume that the radius of this planet, R_{Planet} is $3^{1/3}$ times greater than that of the Earth, or 9.2×10^8 cm. If the planet is in a circular orbit at the most favorable inclination of 90°, it would produce a total amplitude of radial velocity variation of the M dwarf of 17 m s^{-1}, which is easily measurable with current techniques.

If p denotes the geometric albedo of the planet, then the ratio ε of the observed flux from the planet to that of the star is:

$$\varepsilon = p\,(R_{Planet}/D)^2 \qquad (2\text{-}2)$$

The value of p depends on the amplitude and angular dependence of the various sources of scattering in the planetary atmosphere, integrated over the surface of the sphere (Charbonneau, et al., 1999). For a Lambert-law sphere, $p = 2/3$. Therefore, in this example, $\varepsilon = 5 \times 10^{-3}$. Charbonneau, et al., (1999) already have been able to measure a contrast between π Boo and its companion at this level ($\varepsilon = 5 \times 10^{-3}$) with Keck; therefore, a similar spectroscopic investigation of terrestrial planets with CELT is realistic.

There are other possible methods for performing spectroscopic studies of terrestrial planets. If the planet happens to eclipse the central star, as occurs with HD 209548 and its companion, then absorption lines produced in the atmosphere of the planet may be studied during the eclipse. Also, the duration and amplitude of the eclipse will allow a direct determination of the radius and albedo of the planet.

Spectroscopy in the mid-IR should be an important tool to study the planet's atmosphere. If the planet is at a distance of 10^{11} cm from the star, which is only a factor of 2.5 greater than the distance between the Earth and the Moon, it is likely that the planet's rotational period will be tidally locked to its orbital period. This means that the planet will always present the same face to the star it orbits. Consequently, although there will be a range of temperatures on the surface, the mean temperature on the illuminated side of the planet, T_p, if the albedo, w, is 0.3 will be given by the expression:

$$T_p^4 = (1 - w)\,L_*/(4\,p\,s_{SB}\,D^2) \qquad (2\text{-}3)$$

where s_{SB} is the Stephan-Boltzmann constant. With the parameters given above, then $T_p = 1000$ K.

Therefore, if we assume that both the planet and the star radiate like black bodies, the ratio of the thermal flux from the planet to that from the star is given by the expression:

$$F_\nu(\text{planet}) / F_\nu(*) = (R_p / R_*)^2\,(B_\nu[T_p] / B_\nu[T_*]) \qquad (2\text{-}4)$$

At 10 μm, $F_\nu(\text{planet}) / F_\nu(*) = 9 \times 10^{-4}$. With an S/N of about 1000, it should be possible to measure the spectrum of the planet since in the mid-IR we expect that the spectrum of the planet and that of the central star are very different from each other. Furthermore, the planet's spectral lines will exhibit strong, periodic, predictable Doppler shifts that will provide an additional signature of light from the planet.

While we do not imagine life similar to that on the Earth to exist at 1000 K, the dark side of the planet will be much cooler. As a result, there may be a habitable zone on the surface of the planet.

One of the spectroscopic signatures of life on a planet might be the presence of O_2. This molecule is very difficult to measure from the ground. However, its daughter molecules, OH and O_3, may also be indirect signatures of the presence of biological processes in the atmosphere of an extrasolar terrestrial planet. We should be able to detect OH in the near-IR and O_3 in the mid-IR in the spectrum of such a planet. That is, although these two molecules are present in the Earth's atmosphere, they would exhibit

a predictable velocity shift of 140 km s^{-1} with a period of ~ 0.5 days in the spectrum of the extrasolar planet. This predictable shift will enable us to separate the telluric lines from those intrinsic to the planet, and we will be able to measure the amount of OH and O_3 in the planet's atmosphere. These molecules may signal that life exists on the planet. With CELT, we will be able to begin a focused scientific search for signs of life elsewhere in our Galaxy.

Direct Imaging of Extrasolar Planets

A powerful complement to the indirect planet-detection techniques (radial-velocity measurements and astrometric measurements) and spectroscopic detection discussed above will be *direct* imaging of extrasolar planets themselves by resolving them from the parent star. Such detections would then allow photometry or spectroscopy of detected companions, allowing measurements of their composition and perhaps temperature, and allowing (for example) the rocky giant planets to be distinguished from small gaseous planets. It will also be an important step towards the direct imaging of true Earth analogs, as proposed by NASA's Terrestrial Planet Finder program.

Direct detection of extrasolar planets is extremely difficult. Jupiter is approximately a billion times fainter than the sun. Seen through the Earth's atmosphere, stars are surrounded by a diffuse halo of scattered light. The key to direct detection is to enhance the contrast of the planet relative to this halo. There are two regimes in which this will be possible with CELT: searches for young extrasolar planets through near-IR emission with a normal AO system, and searches for solar systems like our own using high-contrast "extreme" adaptive optics.

Direct Detection of Young Extrasolar Planets

The first regime for CELT direct planet detection is the search for *young* extrasolar planets; at an age of 10 million years, a Jupiter-mass planet would still have an effective temperature of 600-800 K, and be only a factor of 10^3 dimmer than a sun-like star in the near-infrared. Current 8-10-m telescopes with AO systems could detect such objects at separations of ~0.5-1.0 arsecs. Since the nearest populations of young stars (e.g, the TW Hydrae association) are ~ 50 pc from the Earth, this corresponds to a scale of 25 AU; it is currently unknown whether giant planets are common on such wide scales.

In this regime, the adaptive optics system does little to suppress the halo of scattered light, and sensitivity comes primarily from concentrating the light from the planet into a diffraction-limited spike. Based on current Keck AO performance and predictions of CELT AO performance we could therefore expect to detect young planets at separations of ~0.2-0.3 arcsec, corresponding to a scale of 10–15 AU. This opens up the possibility of seeing solar systems like our own in the process of formation, providing a direct test of the conventional planet-formation paradigm.

Detection of Extrasolar Planets in Reflected Starlight

The second regime for planet detection is a search for reflected starlight from Jupiter-like planets orbiting nearby stars. As mentioned above, conventional AO systems with sub-aperture size d = 50 cm have little effect on the scattered light halo. By the time CELT is a reality it will be possible to construct so-called "extreme" adaptive optics (EAO) systems with d = 5-1 cm, using new technologies such as MEMS deformable mirrors. Section 9.5.1 discusses the design and performance of such systems in more detail. Briefly, such systems massively suppress the scattered light halo to an intensity 10^6-10^7 of the central star, and with long (1-4 hr) integration times, could overcome noise from residual halo fluctuations and see Jupiter-like planets at 1-10 AU separations.

Deployed on an 8-10-m telescope, such a system would require stars brighter than $m_R \sim 3.5$, limiting it to a handful of nearby stars. On CELT, such a system would operate to $m_R \sim 5$-6, opening up ten times as many target stars and allowing for a large-scale survey, e.g., of all sun-like stars within 10 pc. As indirect techniques become sensitive to planets in wider orbits, this will also produce several cases in which planets detected by astrometric motions or radial-velocity variations are within reach of direct CELT AO imaging, an extremely powerful combination. It is even possible that around the nearest sun-like stars CELT could achieve contrast levels of $\sim 4 \times 10^{-10}$ at separations of 1-2 AU, sufficient to detect an Earth-like planet, thus paving the way for space-based spectroscopic follow-up.

Technical and Instrumental Requirements

The indirect planet searches described above require a high-resolution ($R \sim 40{,}000$) optical spectrograph, similar to the HIRES instruments on Keck. Spectral multiplexing is not important for this particular application.

The spectroscopic detection of planets requires diffraction-limited intermediate-to-high-resolution spectroscopy in both the near-IR (1-2.5 μm) and in the thermal-IR (5-12 μm).

Direct imaging of young extrasolar planets requires only the basic CELT AO system as discussed in Chapter 9, operated in either laser guide star or natural guide star mode, combined with an infrared camera similar to Keck's NIRC2, with a Lyot mask or other coronagraph. (Segment aberrations, if severe, may require a more sophisticated coronagraph.)

Detection of mature extrasolar planets orbiting nearby stars requires a full-fledged EAO system with 10^3 to 10^6 actuators; this daunting but exciting prospect is discussed in detail in Section 9.5.1.

The two latter projects place requirements on the telescope and site choice. The largest known populations of nearby young stars, such as the TW Hydrae association and Tucanae association, are located in the southern hemisphere from ~25 to ~70 degrees DEC. Although some young associations are being discovered further north, it appears that (for currently unknown reasons) the bulk of the nearby young stellar groups are in the south.

For the EAO mode, by contrast, the primary site consideration is seeing; EAO performance drops sharply with increasing r_0. An EAO system should be located at a site that experiences a significant number of nights with $r_0 > 50$cm in the I-band.

2.5.3 CELT and Star Formation

The development of the theory of stellar structure and evolution is one of the great achievements of twentieth century science. Yet this elegant theory that explains the life cycle of stars is incomplete in one critical aspect: It does not predict nor account for the formation of stars. Star formation plays a key role, at small scales, in the origin of our own Solar System; and at much larger scales, in the appearance, structure, and evolution of galaxies. However, it is the least understood aspect of these fundamental processes. Nonetheless, over the last quarter-century impressive advances in our understanding of star formation have resulted from the continued development of new technological observation capabilities from both the ground and space. During this period we have learned:

1. Stars form continually in our galaxy within the dense cores of giant molecular clouds.

2. The process of star formation is almost always accompanied by the formation of a circumstellar disk. By analogy with the Solar System this suggests that conditions suitable for planet formation may be a natural by-product of the star formation process and that planetary systems may be common in the galaxy. The recent detection of extrasolar planetary systems around a few nearby stars has provided support for that notion.

3. Star formation is a complex and dynamic process dominated by gravitational collapse, but is always accompanied by (and may even require) the energetic ejection of spectacular bipolar jets and outflows.

4. Stars tend to form in pairs, groups and clusters, but rarely in isolation.

Existing theories cannot simultaneously account for all these facts. Moreover, a number of additional mysteries still need to be solved before a credible theory of star formation can be constructed. Perhaps two of the most critical of these issues are: 1) identifying the physical factors that determine stellar mass, and 2) determining the origin of the initial mass function (IMF). Until these issues are resolved our knowledge of the entire life cycle of stars will be incomplete and our understanding of galaxies will remain on a shaky foundation. The physical process of star formation spans an enormous range in both spatial scale (~ 8 orders of magnitude) and density (~ 20 orders of magnitude), and although much has been learned in the last two decades, direct observations of various key stages has proved to be a formidable challenge. In particular, we have little knowledge of the critical processes that occur on relatively small physical scales (< 200 AU), such as the development of energetic bipolar jets, the growth and evolution of a protostar through accretion and infall of circumstellar matter, and the evolution of a circumstellar disk to form a system of planets.

CELT, working in concert with the NGST and ALMA, will have a profound impact in the upcoming decades on our understanding of the origins of both stars and planetary systems. Working at differing wavelengths and probing a range of angular scales (from 1-200 AU), these new facilities will provide a more detailed and comprehensive picture of the earliest stages of star and planet formation than has been previously possible. In particular, the angular resolution and sensitivity afforded by a diffraction-limited 30-m telescope such as CELT provides a unique opportunity to obtain spatially resolved observations of regions as small as 1 AU (at 1 μm) in the nearest protostellar clouds.

Observations with such a large telescope will yield information on:

- **The origin and nature of bipolar jets.** High angular and spectral resolution observations should be able to determine how close to the central protostar the jets are collimated and whether jets form as disk winds or are driven from close to the surface of the protostar itself. Detailed knowledge of the driving mechanism of such jets may be needed to assess whether such ejections regulate the mass of the star and the form of the IMF in the star formation process.

- **The structure and nature of protostar.** High-resolution spectroscopy at near-infrared wavelengths would probe the velocity/density structure of protostellar environments on scales from a few AU down to the stellar surface (even in seeing-limited mode). Protostars gain mass through infall and disk accretion; disk accretion is believed to dominate in the inner regions. However, the nature of the accretion mechanism is unclear. Does material accrete directly from the disk onto the stellar surface or instead along dipole field lines from a truncated disk? Is the accretion steady or episodic?

Is the accretion path the same in protostars as in pre-main sequence stars? Detailed observations by CELT of the inner regions of the protostellar disk and envelope are essential to answering these questions. In addition, the added sensitivity provided by the increased light-gathering power of CELT will enable measurement of the photospheric absorption lines from protostellar atmospheres that are too heavily veiled to be easily detected with smaller telescopes, even with high spectral resolution. This will enable direct measurement of such physical properties as the effective temperatures, surface gravities, rotation rates, and even accretion energy of protostars, and will critically constrain protostellar theory.

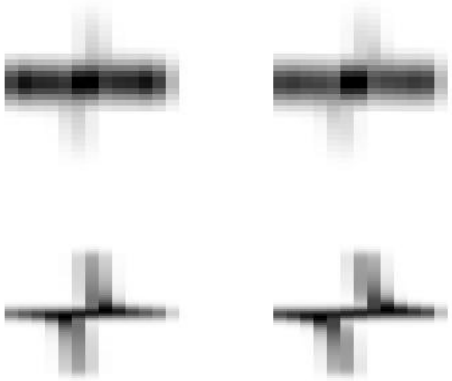


Figure 2-13. The dynamics of both binary stars and disks can be used to measure the masses of young stars, and thereby calibrate the pre-main sequence evolutionary models that currently produce widely discrepant results towards lower masses. The figure shows a simulation (provided by M. Brown) of H2 emission (at 17 μm) from a disk with a 0.5 M_{Sun} star (left) compared to the view of an 0.8 M_{Sun} star (right). In all cases the star is located at 50 pc. In the image, wavelength runs left to right, distance up and down. The full extent of the visible disk is 100 AU in radius with 12 AU per pixel. The velocity scale has 1.5 km/s per pixel. CELT (bottom) compared to Keck (top) observations clearly show how the factor of 3 resolution increase is key. By observing regions closer to the central star we will observe regions of higher velocities, and thus will more easily and more accurately measure rotation speeds.

- **Protostellar companions and masses.** High angular resolution imaging and spectroscopy will permit the measurement of the frequency, separations, and orbital motions of binary companions to protostars and more evolved young stellar objects (such as T Tauri stars) on scales of 1-5 AU in the nearest star forming regions. This would yield the first direct determinations of protostellar masses, knowledge of which is fundamental to the development of a complete theory of protostellar formation and evolution (see Figure 2-13). Moreover, determination of the frequency of protostellar companions is vital to understanding the process of star formation and the survivability of protoplanetary disks.

- **Disk structure and chemistry.** CELT will provide both the spatial and spectral resolution needed to investigate the physical and chemical structure of disks. For instance, the majority of the mid-infrared emission from a protoplanetary disk is confined to the inner circumstellar regions ($r < 20$ AU). The improvement in angular resolution with CELT will allow the first spatially-resolved mapping of the dust structure and chemistry of young disks in the region where planetary systems are thought to form (see Figure 2-14). For both these disks and the older debris disks (for which

the first spatially resolved images have recently been obtained, Figure 2-15), maps of the thermal emission at mid-infrared wavelengths, or scattered light at near-infrared wavelengths, have the potential to reveal gaps and spiral arms in the surface density caused by gravitational interaction with embedded protoplanets. For instance, Jupiter would have first cleared the primitive solar nebula to form a gap of ~1 AU at an orbital radius of 5 AU. This would be detected at the distance to the nearest star formation regions (d = 150 pc).

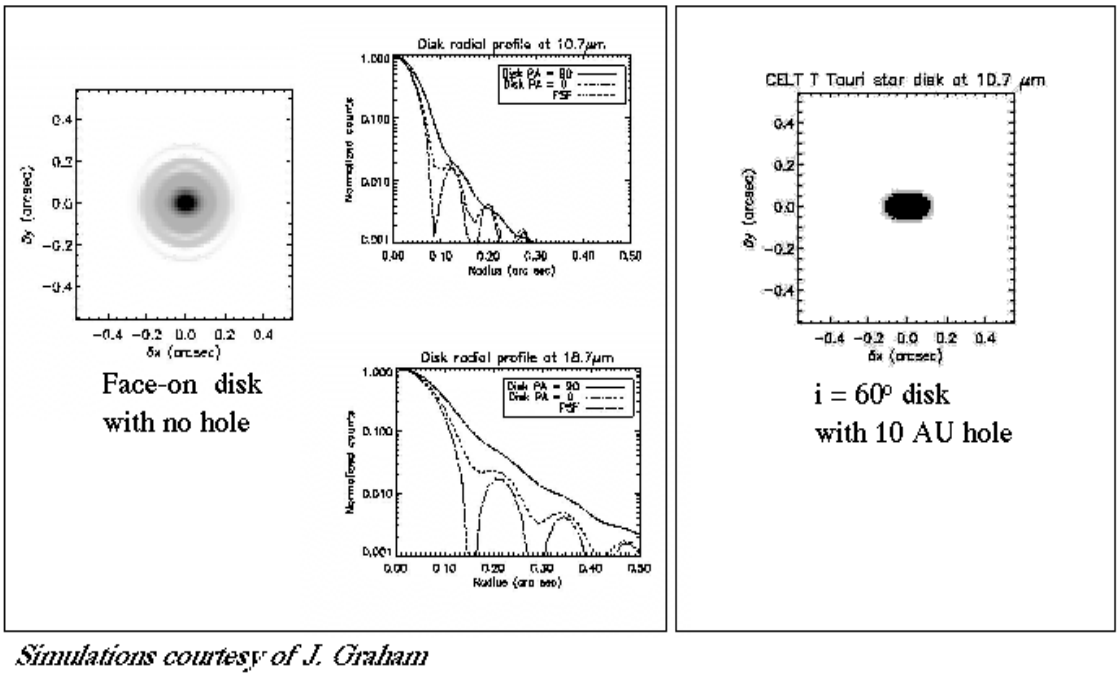


Figure 2-14. Simulations of mid-infrared observations of a disk surrounding a young low mass (T Tauri) star. The disk extent is clearly resolved and detailed structure such as a 10 AU gap can also be detected.

Technical and instrumental requirements

The instrumental requirements for star formation studies with CELT fall into two categories: high-order AO-based imaging and spectroscopic observations in the 1-5 μm range, and diffraction-limited imaging and spectroscopy in the thermal-IR (5-30 μm). In the near-IR a relatively narrow field, high Strehl AO imaging system, and the ability to do spatially-resolved spectroscopy (e.g., using an IFU) would be ideal. It is not foreseen that a patrolling multi-headed IFU system, or a particularly wide AO-corrected field, would be essential for most star formation science.

The requirement for the thermal-IR observations has the potential for acting as a much stronger driver for aspects of the telescope and low-order AO system design, as well as for the choice of the CELT site. With the exception of the 10 μm atmospheric window, which is quite transparent (see Section 2.3.2 above), the thermal-IR transmission depends critically on the water vapor content in the atmosphere. The best terrestrial sites average 1 mm of precipitable water vapor or less (e.g., Mauna Kea and sites in the high Andes in Chile); other sites that have been developed recently (e.g., Cerro Paranal in Chile) have much worse statistics. The water vapor content is strongly correlated with the altitude of the

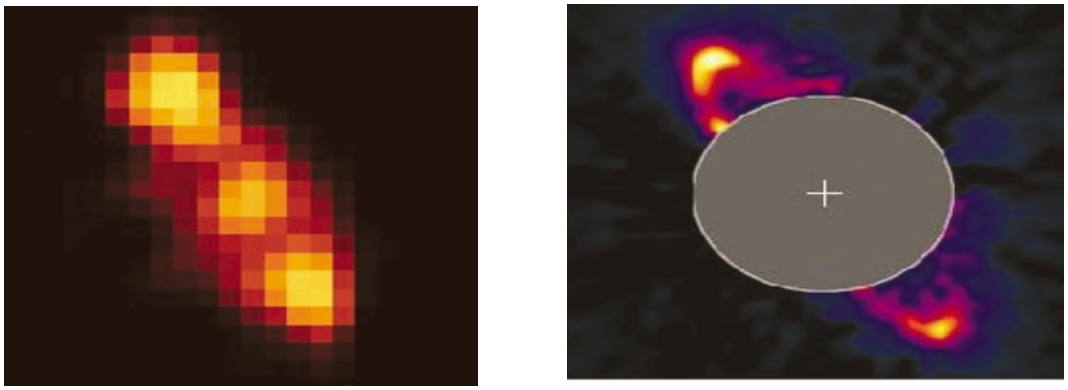


Figure 2-15. Direct images of the debris disk around the main sequence A star HR 4796, which show both the advantage of high spatial resolution and the interplay between ground- and space-based facilities. On the left is a mid-infrared (24.5 μm) image from the Keck Telescope of the thermal emission from the disk. On the right is a near-infrared (1.1 μm) image taken with NICMOS aboard Hubble Space Telescope, which detects the disk in light scattered from the central star. With CELT, much higher resolution images could be obtained of these disks around A stars, and of disks around lower mass and younger stars that are currently unobtainable.

observatory. As discussed above in Section 2.4.2**,** the thermal IR observations with CELT cannot compete in raw sensitivity with NGST, but all of the science discussed above is driven by the five times higher spatial resolution afforded by CELT, and by the ability to obtain very high spectral resolution mid-IR spectroscopy.

Two different mid-IR instruments are suggested: a 5-30 μm imager, and a 5-30 μm spectrograph capable of R ~ 100,000. As discussed in Section 9.5, there are several concepts being explored for a mid-IR optimized AO system and focal position. The emissivity of the telescope combined with the AO system may be prohibitive if the same AO system is used in the thermal-IR as in the near-IR.

2.5.4 Nearby Galaxies: Chemical Evolution and Star Formation Histories

Introduction

The vast majority of galaxies are studied in the merged light of millions or billions of stars and other glowing gases. In our own Galaxy, on the other hand, we can study individual stars and have been able to construct a remarkably clear picture of the star formation history, the gradual buildup of elements higher in atomic number than helium, and the kinematic and dynamical processes that shape the Galaxy. There are still a number of unresolved issues in our understanding of the history of the Galaxy. One such issue is the relative contribution of a global initial collapse of gas and dust versus the incorporation of dwarf galaxies or galaxy fragments through tidal interactions (as predicted by hierarchical structure formation). Because stars with initial mass slightly lower than the Sun have lifetimes comparable to or greater than the age of the Galaxy, we have many examples of stars that formed in the initial collapse of the Galaxy as well as stars formed throughout the history of the Galaxy. This is what is known as the "fossil record" for Galactic history.

It is already known that the complement of dwarf galaxies around the Galaxy had their first burst of star formation synchronized remarkably well with that of the Galaxy. However, the subsequent star-formation histories and chemical enrichment histories are different in almost every case for the companion galaxies of the Milky Way. With HST and the Keck 10-m telescopes the first steps have been made toward

detailed study of the fossil record of old stars in M31 and some of its companions. One very exciting capability of a 30-m telescope is the extension of detailed fossil record studies to other members of the Local Group and beyond.

Photometry

The color-magnitude diagram studies of star clusters and dwarf galaxy companions of the Galaxy have been used with great success to estimate distances, overall metallicity, and ages. Specific useful observations are:

- apparent level of the horizontal branch (HB) as a standard candle for distance measurements,
- slope and color of the red-giant branch (RGB) along with its intrinsic color width for estimating [Fe/H] and metallicity distributions,
- the apparent magnitude of the main-sequence turnoff (MSTO) for direct estimates of the age(s) of a stellar population, and
- the apparent brightness of the tip of the RGB for estimating distances.

The typical surface density of giants at the center of a dwarf spheroidal galaxy at the distance of Virgo is around 1000 per square arcsec. For seeing-limited observations this density makes observations of individual stars impossible. However, at the diffraction limit of a 30-m telescope this is less than 0.05 stars/resolution element at R. The local universe out to Virgo contains hundreds of dwarf galaxies, and many massive spiral and elliptical galaxies. The sample of objects for which we would have accurate estimates of star formation histories and chemical enrichment histories would go from the handful in the Galaxy's complement of dwarfs to hundreds, spanning a huge range in environment. Important issues that could be addressed are: the importance of local galaxy density in determining star formation history; the dispersion in the chemical evolution histories for galaxies of similar mass; and the very puzzling situation that galaxies apparently similar in total mass and environment can have very different histories of using and losing their initial gas content.

Moderate- and high-resolution spectroscopy

The chemical evolution history of a galaxy, or component of a galaxy is written in the details of the distribution of elements seen in the atmospheres of stars. Particularly with Keck high-resolution spectroscopy we are beginning to piece together the details of the chemical evolution of the Galactic halo going all the way back to the first epoch of star formation. The principal inputs are relative distributions of elements with different nucleosynthetic origins. With high-resolution spectroscopy limited to relatively bright stars with 10-m telescopes, the nearby galaxies are so far unexplored.

It is in the realm of $R > 5000$ spectroscopy that CELT will excel. There are a large number of open questions in Galactic stellar astronomy that have proven too difficult for the current 8-10-m telescopes. Extending abundance and kinematic studies for individual stars to other Local Group galaxies and beyond will be a very important strength of CELT. $R \sim 5000$ spectroscopy allows absorption-line velocities to be determined with a precision of 1 km/sec, and abundances of certain elements to be measured based on equivalent-width measures for individual transitions. For the I through H bands, this resolution is sufficiently high to resolve the OH emission of the sky and allow work in uncontaminated regions of the spectrum (see Figure 2-6).

Higher spectral resolutions (typically 25,000 < R < 60,000) are required for detailed chemical abundance studies that include elements throughout the periodic table and represent all of the nucleosynthesis paths in stars and supernovae explosions.

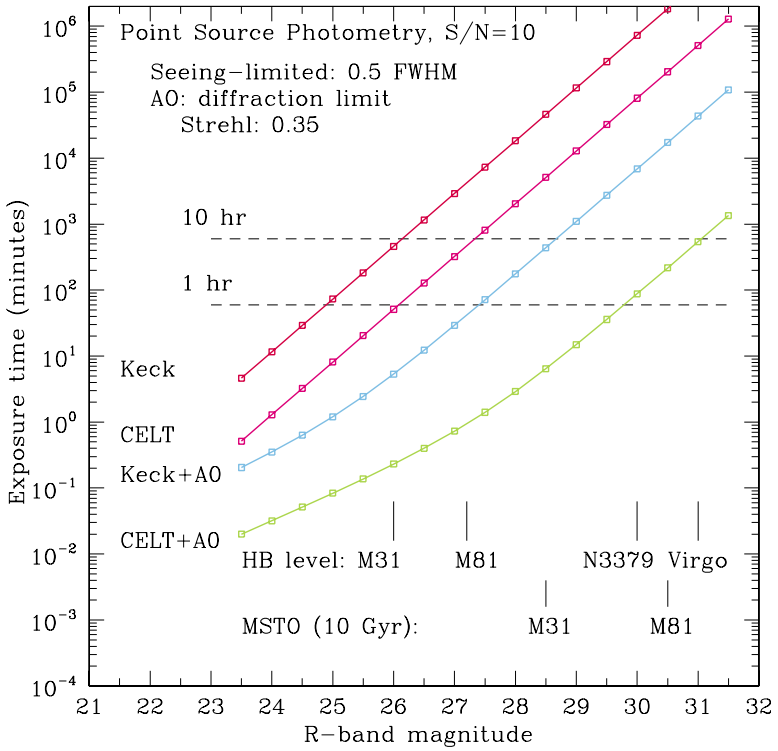


Figure 2-16. The integration time required for R-band imaging as a function of magnitude to reach S/N=10 for a 10-m telescope and 30-m telescope operating in seeing-limited and diffraction-limited modes. For AO correction at this wavelength a Strehl of 0.35 is assumed. Also shown is the level of the HB and an old population main sequence, as they would appear at the distance of M31, the M81 Group (4 Mpc), NGC 3379 in the Leo Group (10 Mpc) and at the Virgo Cluster (16 Mpc). For AO observations with CELT, RGB stars are relatively easily observed throughout the local universe to the distance of the Virgo cluster. The horizontal branch is accessible with some effort with CELT and AO in this same volume. Direct measurements of an old population's MSTO will be possible throughout the Local Group and at the distance of the nearest neighbor groups around M81, NGC 5128 and Cen A.

State-of-the-art observations of these kinds with 8-10 m facilities are limited to the nearest members of the Local Group (out to M31) for R ~ 5000 studies (kinematics and rough chemical abundances) of bright giants. Detailed abundance studies at higher spectral resolution are limited to V < 16 – giants in the halo and dwarf stars in the solar neighborhood only.

There are tremendous possibilities in these areas with CELT both in extending work in the Galaxy and moving out into the Local Group and for some projects into the nearby M81 (north) or Cen A (south) groups.

Specific areas that will have to await a 30-m telescope for further progress are:

- in situ detailed abundances for Galactic main-sequence globular cluster stars;
- the properties and number density of halo white dwarfs (currently thought to be a significant contributor to the Galactic dark matter halo based on the MACHO project);
- detailed abundance studies for giant stars in the outer Galactic halo, M31, the M31 dwarf companions, and M33;
- detailed abundance studies of early-type stars in low-metallicity environments (dwarf galaxies).

Technical and instrumental considerations

Direct imaging color-magnitude diagram studies require reasonably high Strehl observations (> 0.3) to 800 nm over modest fields (10-30 arcsec diameter). Traditional stellar abundance studies have been carried out shortward of 600 nm where there is an abundance of atomic absorption lines. Rarely are spectroscopy studies in this regime background-limited, and AO-fed spectroscopy does not result in significant gains. Much of the chemical abundance work described above would be carried out in the seeing-limited mode.

2.5.5 Probing Galactic Nuclei with CELT

Super-massive black holes are believed to provide the central engines that power the abnormally luminous galaxies called quasars (with luminosity exceeding 10 times that of a normal bright galaxy), and the somewhat less luminous, but still anomalously bright, galaxies called "active galactic nuclei" or AGNs. This insight was achieved more than two decades ago.

We are now learning that super-massive black holes are common at the centers of nearby galaxies, with masses ranging from 10^6 to 10^9 M_{Sun}. After years of work combining HST imaging with HST and ground-based spectroscopy, an incomplete census of ~35 super-massive black holes among the nearest galaxies has been assembled. The progress of this work was fueled by the recent installation of STIS on the HST, which offered long-slit spectroscopy with a narrow slit, thus enabling astronomers to take full advantage of the superb spatial resolution of HST in the spectroscopic mode.

As shown by Ferrarese and Merritt (2000), and also by Gebhardt, et al., (2000), there is a close relationship between the deduced mass of the black hole and the velocity dispersion of the host galaxy's bulge. A summary of the current state is shown in Figure 2-17, and a prediction of M_{BH} is given, with surprisingly small scatter, by

$$M_{BH} = 1.2 \times 10^8 M_{Sun} \ (\sigma/200 \text{ km s}^{-1})^{-} \qquad (2\text{-}5)$$

where σ is the bulge velocity dispersion. This implies that the presence and ultimate size of the central black hole must be closely related to how the galaxy itself formed. A correlation with somewhat larger scatter between the black hole mass and the bulge luminosity of the galaxy,

$$M_{BH} = 0.9 \times 10^8 M_{Sun} \ [L_B \text{ (bulge)} / 10^{10} L_{BSun}] \qquad (2\text{-}6)$$

has been known for the past few years (Kormendy 2000). These relationships apply to galactic bulges, but *not* to galactic disks.

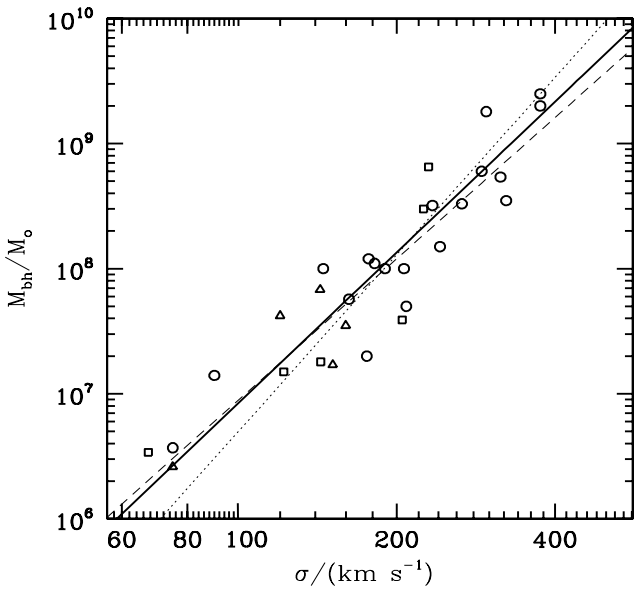


Figure 2-17. The empirical relationship between inferred black hole mass and galaxy central velocity dispersion (Gebhardt, et al., 2000).

This new insight suggests that most galaxies contain super-massive central black holes, but at the present time these are relatively inert and not being "fed" enough gas to produce high-luminosity central sources. This paradigm offers new possibilities for constraining the process of galaxy collapse and formation, and understanding the dynamics of the central regions of galaxies. Theoretical efforts are already underway to understand the relationship between central black holes and star formation rates, gas dynamics and other forms of feedback (Ciotti and Ostriker 1997, 2000; Blandford 1999; Silk and Rees 1998). We also want to understand the complex issues associated with transferring angular momentum outwards so that the central engine can be "fed" by infalling material from a massive accretion disk.

We believe that understanding the phenomena associated with super-massive black holes in the nearby universe, and exploring this phenomena in more distant galaxies are significant tasks for which CELT science will bring major advances. Kinematic estimates of the masses of central black holes in AGNs rely on either measurements of the rotation curves of gas circulating around the black hole or measurements of the velocity dispersion of the stars in its close vicinity. In either case, the high spatial resolution of CELT will be critical to such studies.

AGN Accretion Disk Sizes and Spatial Resolution Issues

In the case of QSOs and Seyfert nuclei, the central black hole is fueled by an accretion disk that is responsible for thermal radiation at a temperature of around 20,000 K, and produces the "big blue bump" at around 2500 Å on top of the power law continuum due to synchrotron radiation. The size of the optical disk is about 10^{14} cm for a typical Seyfert galaxy and 10^{16} cm for a QSO. The broad emission lines characteristic of QSOs and Seyfert galaxies are produced in a region about ten times bigger than the optical disk, although the relationship between the two phenomena is not understood.

These structures immediately surrounding the central engine and responsible for its radiation cannot be resolved by even a 30-m telescope.

Black hole masses, on the other hand, can be determined using a number of techniques. The first is the dynamical perturbation introduced by the central super-massive object. The presence of a central black hole is deduced from a rise in the velocity dispersion of the stars as the central point mass begins to dominate the gravitational potential. For a black hole of mass M_{BH} in a galaxy with a stellar velocity dispersion of $\sigma_V = 300$, this happens at a radius r_D such that $GM_{BH} / r_D \sim \sigma_V^2$ or $r_D \sim 6$ pc for $M_{BH} = 3 \times 10^9 M_{Sun}$. The perturbation to the central velocity dispersion can be measured out to radii $r \sim 10\ r_D$ (depending on the surface brightness of the stellar bulge), or perhaps ~ 60 pc (about 0.6 arcsec at the distance of the Virgo cluster). This scale is so small that for higher redshifts or smaller black hole/ galaxy masses, and given the preponderance of nuclear star formation at earlier cosmic epochs, most of the measurements are likely to be based on gas dynamics of the nuclear disks rather than on traditional measures of stellar velocity dispersion.

HST observations have shown that many local AGNs have a relatively large rotating disk of emitting gas which is responsible for spatially unresolved sharp emission lines observed in ground-based spectra. As an example, the emission line disk seen in M87 by HST is about 150 pc in diameter and is believed to feed a black hole with mass $3 \times 10^9 M_{Sun}$. (The black hole mass deduced from the rotation of the accretion disk gas agrees with that obtained from observations of the velocity dispersion of the stars in the vicinity of the center of M 87.) A recent study by Sarzi, et al., (2000) of 24 nearby AGNs using HST/STIS revealed evidence for emitting gas in the majority of the objects. In a substantial minority the velocity field was sufficiently symmetric to be attributed to rotation and the mass of the central black hole could be estimated. These emitting disks range in size from a few pc to as large as ~ 150 pc for the most massive black holes.

The spatial resolution of CELT, if diffraction-limited, will be $\theta = \lambda / 30\ m = 3.3 \times 10^{-8} (\lambda / 1\ \mu m)$ radians, which corresponds to 0.007 (λ/ 1 μm) arsecs. At a wavelength of ~ 1 μm, this corresponds to a spatial resolution of 0.5 pc at the distance of the Virgo cluster, and 35 pc at $z = 0.5$ (at which point the Hα line is redshifted into the 1 μm window). Thus, it may result that black hole demographics can be extended, especially for the most massive cases, to cosmological redshifts. Because the physical scale corresponding to a fixed angular resolution changes only very slowly beyond $z \sim 0.5$ (see Section 2.4.6: even at $z \sim 3$, the CELT resolution element at 1 μm corresponds to ~ 60pc), *the CELT diffraction-limited angular resolution may well be high enough to resolve nuclear disks around super-massive black holes to the highest redshifts.*

The range of sizes and rotation speeds of these outer gaseous disks is not known; it is also not known if they exist around the central engines of QSOs. A search for such structures in both active and quiescent galactic nuclei would be a prime subject for CELT. In any case, the high spatial resolution of CELT will allow unprecedented ability to discern non-stellar nuclear activity in galaxies at all redshifts for which they can be detected (perhaps $z \sim 5$).

Technical and instrumental considerations

An integral-field spectrograph operating under diffraction-limited conditions would be ideal for this application. There is no need for multiple movable probes, but rather for coverage by a single IFU over an area as large as possible, and certainly not less than 2 x 2 arcsec. The velocity resolution

required, given the intrinsic stellar velocity dispersion of galactic bulges and the rotation speeds of the disks, is ~ 30 km s^{-1}, for a spectral resolution of R = 10,000.

Direct imaging, both through narrow- and broadband filters, at the full diffraction-limited spatial resolution over a field at least 10 x 10 $arcsec^2$ will also be required in order to help disentangle stellar and nebular components in the galaxy centers.

Future developments in extreme AO (i.e., extending AO capabilities into the optical) have the potential to increase the resolution by a factor of 3-4 over that in the near-IR and could conceivably improve sensitivity to nuclear activity as well, allowing access to nuclear black hole studies for even intermediate mass black holes during the "epoch of galaxy formation" at z ~ 1-4. This capability would provide access to black hole statistics during the very time that the stellar bulges are being formed, allowing a direct assessment of the origins of the correlations seen locally.

2.5.6 Diffraction-Limited Studies of the History of Galaxies: The z = 1-5 Universe

It is anticipated that a major activity of the next decade for the current 8-10-m telescopes will be surveys and global statistics of galaxies and QSOs as a function of look-back time, and that substantial progress will have been made by 2010. Placing these objects into a cosmological context, and understanding both obvious and subtle forms of evolution as a function of time, will require delving into the detailed astrophysics of individual galaxies. This would involve measuring small-spatial-scale internal kinematics, chemical abundances, abundance gradients, gas-phase physical conditions, stellar content, sub-kpc morphology, etc., all as a function of large-scale environment and of cosmic time.

There is little question that NGST will in most cases be the preferred facility for observing galaxies at $z >> 5$, where the most important diagnostic features in the spectra of galaxies move into the thermal IR; and for opening up the currently unexplored "dark ages" prior to reionization. Beyond 2.5 μm, the improved spatial resolution from a 30-m ground-based telescope will be nullified by the prohibitive background for faint object science (see Section 2.4). However, for tracing the evolution of galaxy populations over the period of cosmic history most important in forming the stars and metals present in the universe today, more modest redshifts -- but significantly greater spectroscopic capability -- are required. As discussed in Section 2.4, at high spectral resolution ($R \geq 5000$), the terrestrial background in the 0.8-2.5 μm range can be reduced to within a factor of ~10 of that in space, as the bulk of the background comes from very narrow OH airglow lines and not from thermal emission. Coincidentally, this is also the regime in resolution where spectroscopy opens up new possibilities for studies of distant galaxies. These include measuring the relationship between luminosity and mass and measuring the chemical properties of galaxies. These will be made using nebular line diagnostics, the rich interstellar absorption line spectrum in the rest-frame UV, and through the integrated stellar light. These observations will be necessary for understanding galaxy formation and evolution and how they are related to the development of large-scale structure of the universe.

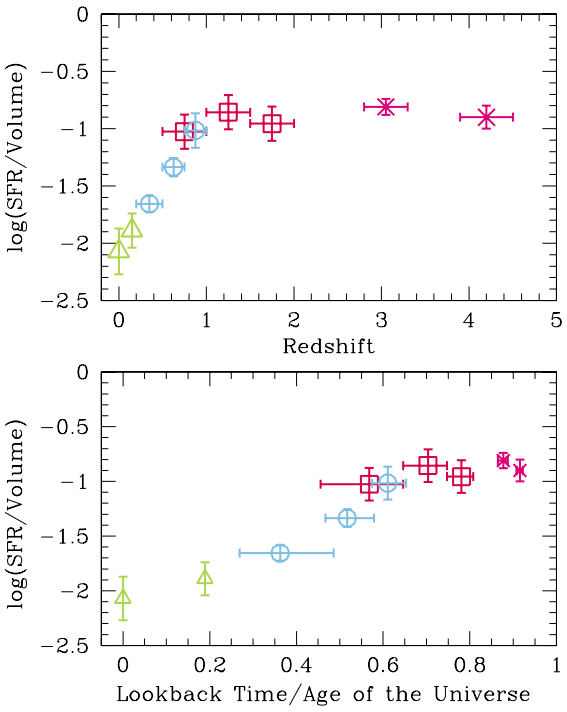


Figure 2-18. Two views of the star formation history of the universe, based on current data. Note in the bottom panel that there is essentially no cosmic time prior to z ~ 5. NGST will explore the very high redshift universe; large surveys for galaxies in the redshift range 1 < z < 5 will be carried out using 4-8-m class survey telescopes and 8-10-m telescopes for follow-up spectroscopy over the next decade. A 30-m, diffraction-limited telescope will provide detailed access to the chemical and dynamical history of the z = 1-5 universe through spectroscopic and imaging capabilities that will be unparalleled.

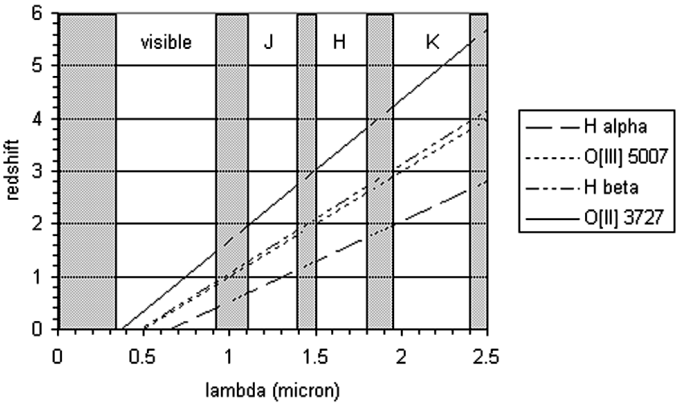


Figure 2-19. Plot showing the accessibility of important diagnostic spectral features as a function of wavelength and of redshift. The nebular emission lines of [OII], [OIII], Hb, and Ha are expected to be important for measuring both chemical abundances and kinematics of galaxies during the z = 1-5 cosmic epoch, and have the advantage of increasing the detectability of galaxies by a factor as large as 5 magnitudes. The brightest objects at these redshifts will be amenable to spatially resolved continuum spectroscopy, capable of yielding information on (e.g.) stellar abundance, age gradients, and stellar velocity dispersions.

Ideally, one needs spectral resolution of $R \geq 5000$ in order to resolve the rotation curve or velocity dispersion of a (potentially) low-mass galaxy. The typical half-light radius of galaxies in this range of redshift is ~ 0.2-0.3 arcsec. We do not know at present if these galaxies are rotationally supported, as resolved spectroscopy is currently difficult or impossible. Adaptive optics on 8-m class telescopes will allow (with a great deal of effort) diffraction-limited images of these distant galaxies that will be significantly better than those achieved so far using the HST; however, they will generally lack the sensitivity to exploit the spatial resolution with spectroscopy. The physical measurements allowed by high-dispersion spectroscopy are key for relating one observed epoch to another, and for connecting theory to observation. Detailed spectroscopy at high spatial resolution will be capable of revealing the physical processes behind the observed morphologies.

With a 30-m diffraction-limited telescope we will be able to achieve the same kind of spatial resolution on a galaxy at z = 1-5 as can presently be achieved at the distance of the Virgo cluster with typical seeing-limited resolution (~ 50 pc per resolution element). This would place as many as ~ 50 resolution elements across a typical compact galaxy at high redshift. These galaxies currently appear as small "fuzzballs" even at HST spatial resolution of ~ 0.1 arcsec. We expect that at much higher spatial resolution they will break up into very small, luminous "knots," making it possible to measure chemical abundances for individual star clusters/giant HII regions, trace the kinematics of large-scale outflows across the face of the galaxies (using the interstellar absorption lines against the UV continuum produced by massive stars), and see whether the overall kinematics of the galaxies are chaotic or follow an underlying ordered motion. It should, therefore, be possible not only to measure the masses of the (generally) compact galaxies observed at high redshift, but also to delve into the detailed baryon physics that controls the appearance and evolution of the galaxies.

The 0.6-2.5 μm range will contain information on the far-UV to the optical/near-IR for galaxies in the redshift range $1 \leq z \leq 5$. This range-of-rest wavelength contains information about the most massive stars, the physics of the interstellar medium, the chemical abundances in HII regions, and the stellar features most commonly used to measure velocity dispersions and age-sensitive line indices for nearby galaxies. The 30-m diffraction-limited telescope will essentially allow *diagnostic* study of galaxies during the epoch of galaxy formation that is equivalent to the current state-of-the-art study of nearby galaxies.

There are several outstanding questions that might be answered. When did galactic bulges form? Are distant galaxies rotationally supported? What controls the decline in the global star formation rate that begins at z ~ 1? What is the mass function (as opposed to the luminosity function) of distant galaxies? How much metal mass is ejected from galaxies during their robust star-forming phase, polluting the IGM? Are chaotic morphologies really indicative of mergers, or are they a natural consequence of rapid star formation? What has been the influence of the large-scale environment on the detailed evolution of galaxies? What controls the epoch when recognizable spiral disks appear? Answers to all of these questions require a combination of spectral and spatial resolution (and the necessary sensitivity when operating in this mode) that is beyond the capabilities of either 8-10-m telescopes *or* NGST.

Practical Issues and Limitations

How will CELT with AO complement and compete with NGST and ALMA, the forefront new facilities that will become operational on roughly the same timescale?

The issues of sensitivity for CELT and NGST are reasonably clear, and are summarized in Section 2.2 of this document. CELT is inferior in sensitivity relative to NGST longward of 2.5 μm, but achieves

superiority for any spectroscopic applications shortward of 2.5 μm. ALMA will be superior to CELT for studying the molecular gas content of high redshift galaxies, and for studying the thermal dust continuum in the observed-frame sub-mm. Here we consider the efficacy of AO-fed observations in the 1-2.5 μm range with CELT.

Galaxies are *not* point sources, and the strength of any general statement one can make is based on the statistics of the sample rather than on any spectacular single observation. Just to set the stage, following is a summary of the spatial resolution of CELT and NGST and the projected physical scales at a variety of redshifts. The assumed cosmology is $\Omega_m = 0.3$, $\Omega_\lambda = 0.7$, $h = 1$.

Table 2-2. Resolution versus Redshift at 2 μm

Redshift (Z)	**Scale** (kpc/arcsec)	**NGST** (diff limit)	**CELT** (diff limit)
0.5	4.3	350pc	70pc
1.0	5.5	450pc	90pc
2.0	5.9	475pc	95pc
3.0	5.4	450pc	90pc
4.0	4.9	400pc	80pc

We already know that the bulk of galaxies in the universe at $z > 1$ have very small physical sizes, with half-light radii on the order of 1-2 kpc. This means that the 0.06 arsec resolution of NGST at 2 μm will barely resolve such objects. With CELT resolution we can place 10-20 resolution elements across a typical compact high redshift galaxy. We do not know what these objects will look like at such high spatial resolution, and the gains with AO will depend strongly on this unknown. We have hints, however, from starburst galaxies in the relatively local universe, that the compact galaxies will become resolved into a number of "super star clusters." The spatial resolution of CELT/AO approaches the physical size of individual HII regions in the local universe.

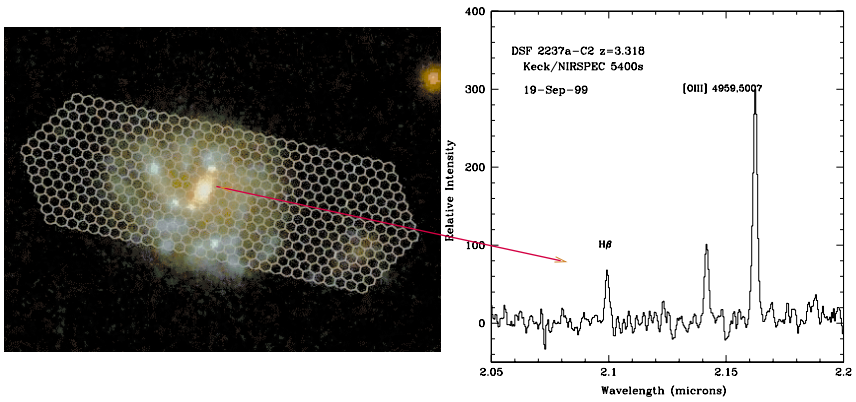


Figure 2-20. The proposed most-efficient mode of observing faint galaxies with CELT will involve integral field spectroscopy, where multiple IFU "units" can be deployed on interesting objects within the AO-corrected field of view. 3-D spectral maps are produced, where each spatial sample within an IFU unit is recorded as a separate spectrum on the detector, as shown schematically in the right hand panel. The quality of the spectrum on the right is roughly that expected for an average 0.05 x 0.05 arcsec:spatial sample on a galaxy at z ~ 3 (although the spectral resolution we are proposing is ~ 3 times higher).

This is interesting from a morphological perspective for galaxies in the redshift range z = 0.5-5 or so; however, the biggest gains of CELT will likely come from (multiplexed) integral field spectroscopy behind AO (see Figure 2-20). The advantages of a ground-based telescope with a large aperture are very clear for spectroscopy, particularly at spectral resolution R ≥ 5000, for several reasons. We have already discussed how high resolution effectively lowers the background by ~ 2 orders of magnitude by resolving out the OH emission. In addition, NGST will most likely not have spectroscopic capability of R >1000. This corresponds to a maximum resolution of ~ 300 km s^{-1}, and while adequate for measuring redshifts of extremely faint galaxies, it is not suitable for detailed kinematics of halos having circular velocities typically in the range 30-200 km s^{-1}. The capability of achieving ~ 80 pc spatial resolution on galaxies in the z = 0.5-5 universe could very well revolutionize the study of very distant galaxies if the resolution is feeding a moderate to high dispersion spectrograph, allowing for spatial dissection into individual knots with accurate velocities and chemistry accessible for each one.

In general, low-resolution (i.e., identification-quality) near-IR spectroscopy of faint objects that are not point sources will be better done with NGST. The CELT/AO limits for faint galaxy spectroscopy will depend very sensitively on what the galaxies look like at 80pc (0.16 arcsec) resolution at 2.2 μm. In the worst case that galaxies are smooth on scales smaller than ~ 0.2 arcsec (the limit of current near-IR images from either ground or NICMOS), then experience with Keck suggests that spectroscopy of objects with K ~ 22.5 should be possible for emission line studies using CELT. Continuum studies might be extended to K~ 20.5 if the observations are background-limited. In the case that our putative galaxy is resolved into many pieces spread over 0.2-0.3 arcsec, it is conceivable that one could reach magnitudes as faint as K ~ 24-25 for successful emission line spectroscopy. This type of observation would be best done with an integral-field-like configuration, where one could focus only on the regions that exceed a certain S/N (unknown a priori). While these limits are still 5-6 magnitudes brighter than the faintest galaxies that deep NGST images will uncover, *they extend to several magnitudes fainter than L* for all redshifts* z = 1–5.

Even at K ~ 22.5 the surface density of galaxies is ~ 50 $arcmin^2$, so that a 1 armin field AO system would include as many as 50 potential spectroscopic targets, and perhaps many more than that. (If one is targeting a particular redshift or range of redshifts then the number will go down significantly, e.g., the number of z = 2-2.5 galaxies in that same collection of 50 would be on the order of 5.) Given the likely limits for spectroscopy with CELT/AO in the near-IR, most of the targets could be easily selected with Keck images, as long as the selection can be done in the optical or near-IR (i.e., no need for NGST). Most of the objects imaged by NGST would be out of the reach of CELT for spectroscopy. We conclude that the CELT AO field size should be driven by science considerations *other than* complementarity with NGST, and even 1 arcmin fields would be scientifically interesting for AO-fed integral field spectrographs. Obviously, larger fields would mean more science per unit integration time.

Note that measuring redshifts for most objects in this same redshift range that have strong spectroscopic features falling in the 1-2.5 μm range can be trivially accomplished using seeing-limited observations in the optical (using CELT, e.g.). The primary purpose of the AO-fed near-IR observations would generally be physical measurements of, e.g., chemistry and kinematics, rather than simply measuring redshifts. In Section 2.5.7 we discuss how optical observations of galaxies at similar redshift will provide complementary data on the physics of massive star formation, galactic winds, and the intergalactic medium; see Section 2.5.8 for a discussion of CELT observations of the "dark ages" prior to z ~ 5.

Instrument Requirements
A summary of the requirements for AO-fed faint galaxy science:

- **Deployable IFUs**: Most of the sky is blank at CELT/AO resolution and there will be perhaps tens of objects per square arcmin within the sensitivity range of CELT spectroscopy. Each IFU need only sample perhaps 1 to 2 arcsec "patches" of sky, feeding a moderately high dispersion R ~ 5000 spectrograph. One does not know a priori how finely one would want to sample spatially on the detector or with the IFU spatial element. In many cases full diffraction-limited resolution may be overkill; depending on the detector characteristics, one might want to have options for spatial sampling to feed the spectrograph.

- **Imaging mode**: An imaging mode covering the 1-2 arcmin AO-corrected field (envisioned to be the maximum in the 1-2 μm range) would be potentially very interesting, but probably less important than the ability to efficiently feed a spectrograph. The "deployable IFU imager" discussed in Section 10.5.3 seems particularly attractive in this regard.

- For faint galaxy science, capabilities beyond 2.5 μm are not deemed important.

2.5.7 Wide Field Science with CELT

The baseline design for CELT provides access to a ~ 20 arcmin field with good images, similar to the Keck telescopes. While at first glance it would seem that the largest gains provided by CELT would be in the near-IR where AO can provide diffraction-limited images, there is a large body of exciting science that takes advantage of the huge gain in spectroscopic throughput (particularly in the 0.3-0.8 μm range where the background is low even at low-to-moderate spectral resolution) afforded by the order-of-magnitude gain in collecting area. If the challenge of building instruments that can take in the whole of the CELT 20 arcmin field can be met, there are huge leaps forward to be made, particularly in our understanding of the evolution of the large-scale structure of the universe.

There is a rich variety of science programs where seeing-limited (or perhaps ground level turbulence-corrected) observations over the full CELT 20 arcmin field will be extremely important. The 30-m aperture of CELT brings objects roughly ten times fainter within reach for moderate-to-high dispersion spectroscopy in the optical; there are many classes of objects whose surface density on the sky is 100 or more times larger at CELT limits compared to (e.g.) Keck. These are the areas where CELT will truly revolutionize wide field spectroscopy. We explore one of these areas below, in order to motivate by example the kind of wide-field instrumental capabilities desirable for CELT.

Galaxy/IGM Connection at High Redshift
By the end of the current decade, we will have very robust constraints on the distribution of galaxies over the redshift range $0 \leq z \leq 1$, largely through ambitious surveys on 4-8-m class telescopes (e.g., the Two Degree Field Redshift Survey, the Sloan Digital Sky Survey, the Keck DEEP survey, the VLT VIRMOS survey, etc.). CELT will enable us to extend this kind of detailed mapping of the universe to the redshift range $2 \leq z \leq 4$, an epoch during which we believe a large fraction of the stars presently seen in galaxies were formed, and where the structures seen at $z < 1$ will be in the early stages of assembly. As we describe below, the insight into the evolution of the entire baryonic component of the universe, and its connection to the underlying distribution of dark matter, may be best attained at these high redshifts, using the unique capabilities of CELT.

Most of the baryons in the universe are believed to be distributed in the form of diffuse gas in the intergalactic medium, a component that has not been observed in its own emission but can be very sensitively observed using background probes in whose spectra the Lyman α line of hydrogen is recorded in the rest-frame UV. Lyman α absorption line studies can detect quantities of neutral hydrogen that are more than seven orders of magnitude smaller than required for detection via emission in the 21 cm line. The Lyman α line is observable from the ground for redshifts $z \geq 1.6$. The undulations in the neutral hydrogen content of the universe along each line of sight to a suitable background source are recorded in the form of a spectrum of the so-called "Lyman α forest." At present, sensitive probes of this dominant intergalactic baryonic component are limited to rare high redshift QSOs, whose surface density on the sky at magnitudes attainable using Keck is very low (much less than 100 per square degree), so that at best only one-dimensional information is accessible. Key to the huge amount of progress that could be made with CELT is that the surface density of suitable background probes depends extremely sensitively on apparent magnitude, and that with a 30-m aperture the number of background targets increases by more than two orders of magnitude. This high density of background probes allows tracing of the three-dimensional distribution of diffuse gas at high redshift. When combined with faint galaxy surveys in the same cosmic volume, which are also enabled with CELT and a 20 arcmin field, a nearly complete census of cosmic baryons, and deep insight into the galaxy formation procåCs and its connection to large-scale structure in the universe, comes within reach.

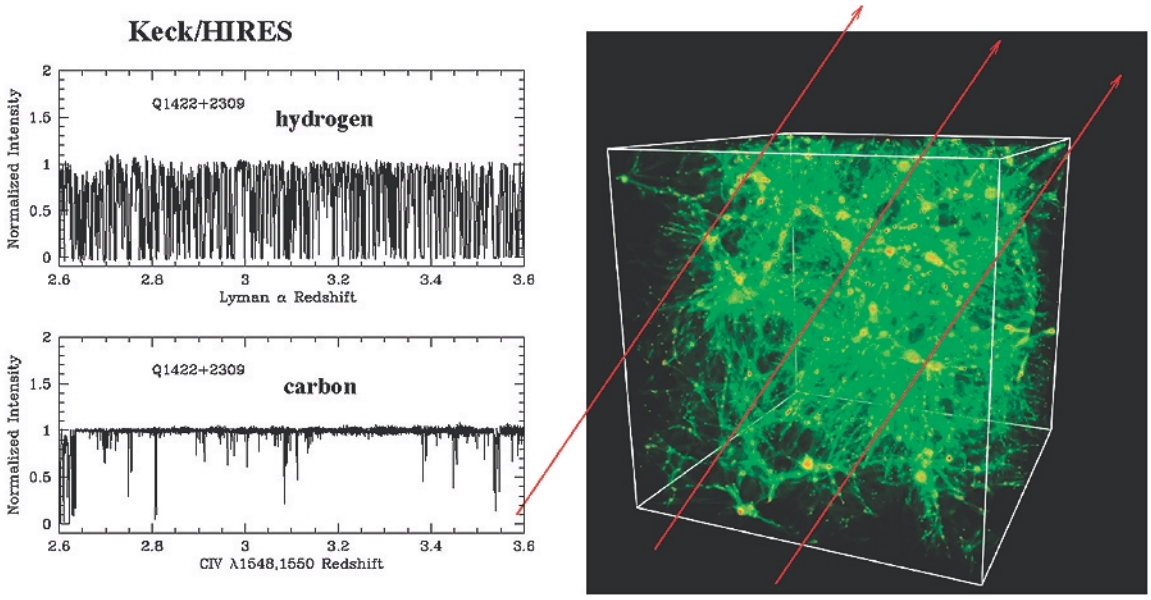


Figure 2-21. An illustration of how multiple lines of sight through a volume of the universe at high redshift (in this case, a hydrodynamic simulation produced by the Princeton cosmology group) can be used to map out the 3-dimensional structure. Each line of sight yields detailed 1-D maps of the H I in the IGM, as well as the associated metal lines, as shown in the panels on the left. CELT allows dense sampling of the IGM because background galaxies can be used, rather than QSOs (which are much rarer).

In Figure 2-21, we sketch a program of observations with CELT that would be capable of surveying both the galaxy distribution and the diffuse intergalactic medium over volumes that are large enough to provide an accurate measure of clustering statistics and characterize the largest structures at $z \sim 2$-3.5. The program would provide a huge range in environment, enabling detailed testing of assertions that the Lyman α forest traces mass fluctuations and that galaxies trace the same fluctuations but in a much

more biased way (the level of bias can in principle be measured directly from these observations, as well). The survey will also allow a detailed map of where the metals are relative to the galaxies, testing ideas about the baseline metallicity in the IGM and where it originated.

The 3-dimensional structure of the IGM has become a focal point for much cosmological theory because the gas giving rise to the Lyman α forest is expected to provide a nearly direct mapping of the total matter distribution (see Figure 2-21). This assertion is based on the application of the so-called "fluctuating Gunn-Peterson approximation," the idea that diffuse baryons in the Lyman α forest trace regions within a factor of ~10 of the mean density of the universe. The "equation of state" of this diffuse H I is simple enough that the H I optical depth should be a monotonic function of the line-of-sight mass density. In this way, "tomography" of the IGM (via multiple lines of sight through any survey volume) should be capable of tracing out the overall mass distribution on all scales larger than about 0.5 Mpc (co-moving). This simple idea of diffuse baryons tracing mass would be extremely powerful, if true. At present, it has not been adequately tested. A clear test of most ideas about galaxy and structure formation would require observations of galaxies and of the relative distribution of the diffuse IGM. These observations are expected to trace the same undulations in the matter distribution, albeit in a "biased" manner. The details of this relationship would constitute crucial constraints on the process of galaxy formation and its connection to large-scale structure. Very recent observations have suggested that the galaxy formation process is so energetic at high redshifts that individual galaxies strongly affect the physical state of the IGM within ~ $500h^{-1}$ kpc (co-moving) through the influence of large-scale winds, which move gas mechanically and shock-heat much of the surrounding medium. This calls into question the simplest assumptions inherent in the current picture of the Lyman α forest; however, the IGM then becomes a powerful tool for understanding the energetics of galaxy formation. In any case, it is clear that the simultaneous study of the diffuse IGM and forming galaxies is fundamental to understanding both.

The same spectra that would be useful for quantitative probes of the IGM distribution would also be of high enough S/N to detect very weak metal lines associated with the same Lyman α forest systems. This would provide probes of metals distribution in the IGM, and allow for detailed chemical and kinematic analyses (of the galaxies themselves) from the high quality rest-frame far UV spectra. Taken together, these observations will provide a three-dimensional mass distribution and a map of the location of gaseous regions that have been enriched in metals.

While the diffuse baryons are probed using intermediate resolution spectroscopy, the details of the galaxy distribution are best obtained through wide-field faint galaxy spectroscopy, capable of obtaining redshifts for objects ~ 10 times fainter than the faint limit for the IGM probes. A low-dispersion optical spectrograph on a 30 m telescope would have the capability of obtaining identification-quality spectra of extremely faint galaxies at z ~ 2-4 (to R ~ 26.5 with high level of completeness (based on scaling of experience from Keck), reaching faint enough to sample the equivalent of the L* galaxy density in the present-day universe. This is a critical aspect of the CELT survey; as it would allow making evolutionary connections with substantially higher validity. Present day galaxy evolution studies, particularly at high redshifts ($z \geq 2$), are hindered by the fact that only relatively rare objects are detectable.

The uncertainties in photometric redshifts are, unfortunately, much too large to allow their use for this part of the project. Accurate redshifts are necessary for examining the three-dimensional galaxy distribution and for establishing galaxies' effects on the IGM. Typical photometric redshift uncertainties at these redshifts would result in distance uncertainties of several hundred Mpc, although they would

clearly be useful in *pre-selecting* the spectroscopic sample. For the sake of illustration, the combination of survey volume and apparent magnitude limits is chosen to approximate the SDSS redshift survey, tuned to the cosmic epoch $2 \leq z \leq 3.5$. There are many reasons for focusing on this redshift range, including the practical ease of obtaining redshifts where the night sky is quite dark (3500-6000 Å), accessibility of the Lyman α line of H I, and the ease of applying photometric pre-selection using *optical* wide-field photometry.

Survey Details

Based on the scaling relative to Keck/ESI, with CELT the apparent magnitude limit for obtaining R = 8000 (seeing-limited) spectra with continuum S/N ≥ 30 per resolution element will be R ~ 24 (for 10-hour integrations). This resolution is high enough that structure in the Lyman α forest is resolved down to velocity scales of ~ 40 km s^{-1}(~500h^{-1} kpc). Inside of this the approximation of the gas as a fluid that closely traces mass fluctuations (the fluctuating Gunn-Peterson approximation described above) must break down due to thermal- and hydro-dynamical effects. While the surface density of QSOs in the required z ~ 2-4 redshift range to R ~ 24 will be about 75 per square degree (estimated from the QSO rate in current Lyman break galaxy surveys reaching similar magnitude levels with high completeness), the corresponding surface density of compact high redshift galaxies will be about 5000 per square degree. While galaxy spectra are somewhat more complex than those of QSOs, it is clear that they can be used as probes since the intrinsic lines are easily separable from the intervening systems because of the large velocity width of the interstellar features in the galaxy spectra. An example of the quality that could be achievable for typical probes is given in Figure 2-22. This surface density places about 1 line-of-sight probe of the IGM at every co-moving square h^{-1} Mpc at z ~ 3, or a spatial sampling rate of one probe per few hundred proper Kpc (and so just about the right spatial resolution to match the spectral resolution).

The survey volume assuming z ~ 2–3.5 will be about 3 x 10^6 Mpc^3 per square degree (Ω_m = 0.3, Ω_λ = 0.7) co-moving. Assuming a faint galaxy magnitude limit of R ~ 26.5, from the known luminosity distribution of UV-selected galaxies at these redshifts, there will be about 50,000 galaxies per square degree in the redshift range of interest. A survey of 10 square degrees would thus contain a volume of a few times 10^7 Mpc^3 (co-moving) and would contain 500,000 galaxies, both numbers very similar to SDSS in the local universe. The sample space density would be ~ 2 x 10^{-2} Mpc^{-3}, or roughly an L^* density in the present universe. The survey should consist of four to six ~ 2.5° x ~ 1° fields, each spanning regions ~ 200-100h^{-1} Mpc (co-moving) transverse to the line of sight, roughly what would be wanted to adequately sample the largest structure that could possibly have existed at those cosmic epochs. There would be hundreds of protoclusters in such a volume.

The point of these numbers is that it would be possible to perform a Sloan-like redshift survey in the $2 \leq z \leq 3.5$ universe, with the added benefit of more than 50,000 "skewers" through the IGM in the same cosmic volume in about one year of observing time with CELT (detailed below). The survey products would include:

- extremely good spectra of 50,000 bright high redshift galaxies (c.f. Figure 2-22), allowing for studies of chemical abundances, detailed kinematics of gas associated with the galaxies, and massive

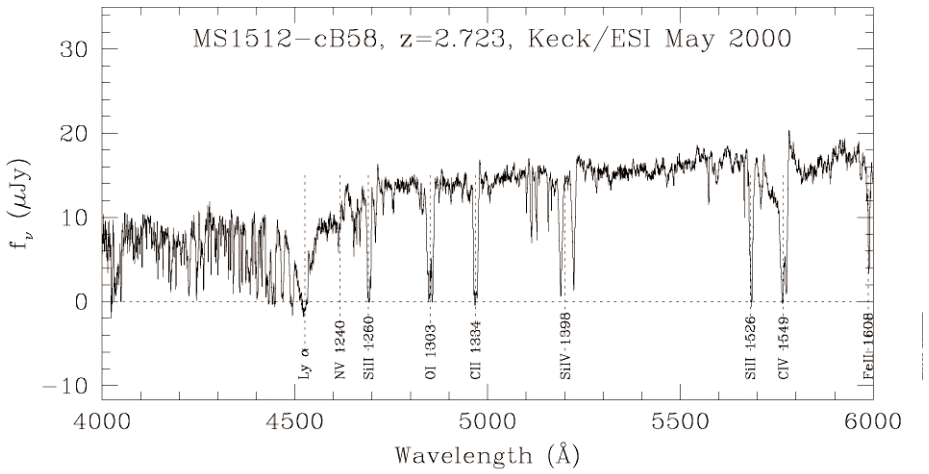


Figure 2-22. Example spectrum of a lensed (by a factor of ~ 30) high redshift galaxy (z = 2.732) obtained using ESI on the Keck II telescope. This is approximately the average quality spectrum that could be expected for R = 24 galaxy probes of the diffuse IGM at high redshift (R ~ 8000). The unmarked doublet features near 5120 Å and 5160 Å are an intervening Mg II doublet at z_{abs} = 0.828 and an intervening C IV doublet at z_{abs} = 2.331, respectively. There are at least 4 additional intervening C IV doublets identified in the spectrum. This spectrum has yielded the gas-phase chemical abundances of 5 different elements and provides our only current glimpse at the detailed physics of high-redshift star forming galaxies. Such spectra will be routine with CELT.

stellar populations from the UV spectral features (this is in addition to the utility of the spectra for tracing out the IGM distribution);

- a densely sampled volume that would contain a snapshot of structure formation progress at only 15% of the universe's current age (the volume is large enough to contain hundreds of protocluster environments, which could be observed conveniently with a small investment of time at other wavelengths because it is contained within fairly small angular scales on the sky, owing to the high redshift);

- constraints on the evolution of galaxy clustering, feedback between galaxies and the IGM, 3-D structure of the IGM, and the relative structures traced by galaxies all with unprecedented precision (compared to *any* redshift); and

- a vast array of follow-up projects using high spatial and spectral resolution, background sources for weak lensing studies at high redshift (z ~ 1-1.5), NGST chemical abundance studies, etc.

Technical and Instrumental Issues

Ideally, a multi-fiber spectrograph with 400-500 fibers over a 20 arcmin field would allow spectra of all of the available R ≤ 24 IGM probes, in the spectrograph field of view, simultaneously. Assuming integration times of 10 hours per pointing, a 10 square degree survey could be completed in 1000 hours (~125 nights). The ideal wavelength range covered would be 3500-7000 Å, with a minimum spectral resolution of R ~ 5000.

The faint galaxy portion of the survey would be slightly more time consuming. Here we assume an *imaging spectrograph* (R ~ 500) using the full 20 arcmin field, with 20 arcmin total slit length and 4 banks of slits across the dispersion direction (so that about 500 spectra could be obtained simultaneously). An integration time of about 2 hours should be sufficient for about 80% spectroscopic completeness and observing 500,000 galaxies would then require 2000 hours, or 250 nights. A less ambitious imaging spectrograph (i.e., a smaller field of view) would require smaller fields, sparser sampling, or both.

2.5.8 CELT and Exploration of the "Dark Ages"

At the time of this writing, very little is known about cosmic epochs prior to $z \sim 4$; there is a handful of known QSOs with $z \geq 5$, and a similar-sized handful of faint galaxies identified, mostly from the Hubble Deep Field. At present, this is one area where the discovery space is still very large, since we do not know at what redshift detectable star formation in galaxies began. The redshifts explored so far do not clearly indicate an absence of star formation well beyond $z \sim 5$. As detailed above, the NGST mission is being optimized for exploration of the $z \geq 5$ universe; however, there are several areas in which CELT may figure prominently for "dark ages" science:

- **Spectroscopy of probes of the re-ionization epoch:** There are currently indications that the highest redshift QSOs are beginning to pierce the re-ionization epoch, when neutral hydrogen in the intergalactic medium was ionized and when the IGM was heated to a temperature of $\sim 10^4$ K. Re-ionization is expected to occur in a "patchy" manner, with initially isolated ionized regions interspersed with regions that remain optically thick in H I. The detailed structure of the IGM requires quite high spectral resolution in order to resolve regions that are optically thin in H I during this transition period. There are estimated to be only 20 $z > 6$ QSOs in the entire Sloan survey, so that clearly a detailed understanding of the physics of re-ionization (and the nature of the ionizing sources responsible for it) will be severely limited. It is likely that many more, much fainter $z \geq 6$ objects (both galaxies and AGN) will be discovered in the intervening decade, but spectroscopic follow-up *at the requisite moderate to high resolution* (R ~ 10,000) will not be possible with either Keck or NGST for these faint objects. CELT may be the only telescope capable of high enough quality spectra to observe the details of re-ionization.

- **The physics of the first galaxies**: Even if the re-ionization epoch is near $z \sim 6$ as we are currently led to believe, the predictions for the formation epoch of the objects responsible for the reionization is not completely clear. Most theory based on hierarchical structure formation predicts that the first objects that can significantly affect the equation of state of the IGM are objects of total mass $\sim 10^8 M_{Sun}$. These would be predicted to have rest-UV luminosities at the nano-Jy level (i.e., $m_{AB} \geq 30$) at $z \sim 10$. Depending on the nature of these sources (sizes surface brightness, spectral features, etc.) and which redshift is most important, CELT's superior performance at the diffraction limit using moderately high spectral resolution (~ 5000) in the 0.0 to 2.5 μm range may be a crucial capability in the NGST era.

- **Narrow-band imaging:** NGST will be limited in what it can do by the filters that go into space with it, necessarily a very finite number. It may well be that narrowband imaging, tuned to (for example) the Lyman α emission line for the highest redshift sources, may reveal interesting physics of reionization. The Lyman α line can be observed up to redshifts $z \sim 19$ using a narrow-band imaging system on CELT constrained to $\lambda < 2.5$ μm.

Technical and Instrument Issues

The area of **spectroscopy of probes of the re-ionization epoch** would benefit from high dispersion, high efficiency spectroscopy in the 0.8-2.5 μm range. It may be that this can be accomplished using the deployable IFUs behind the AO system for the z = 1-5 program. However, it is possible that a normal slit spectrograph optimized for faint near-IR work would be superior, and perhaps even be required to attain the desired spectral resolution of ~ 10,000.

The **physics of the first galaxies** would benefit from near-IR deployable IFU spectroscopy behind AO.

Narrow-band imaging calls for capabilities at optical and near-IR wavelengths. This might be accomplished by a number of different instruments. There is probably a strong science case to be made for "tunable filter" imaging for a wide variety of science applications.

REFERENCES

Blandford, R.D. 1999. "Origin and Evolution of Massive Black Holes in Galactic Nuclei." ed. Merritt, Valluri and Sellwood, in press. Astro-ph/9906025.

Burrows, A., W. B. Hubbard, D. Saumon, and J.I. Lunine. 1993. *Astrophysical Journal* **406**, 158.

Charbonneau, D., R.W. Noyes, S. G. Korzennik, P. Nisenson, S. Jha, S.S. Vogt, and R.I. Kibrick. 1999. *Astrophysical Journal* **522**, L145.

Ciotti, L., and J.P. Ostriker. 1997. *Astrophysical Journal* **487**, L105.

Ciotti, L., and J.P. Ostriker. 2001. *Astrophysical Journal* **551**, 131.

Drilling, A.U., and J.S. Landolt. 1999. *Astrophysical Quantities*, 4th edition, ed. A. N. Cox. Berlin: Springer-Verlag.

Ferrarese, L., and D. Merritt. 2000. *Astrophysical Journal* **539**, L9.

Gebhardt, K., et al. 2000. *Astrophysical Journal* **539**, L13.

Gillette, F., and M. Mountain. 1998. Science with the NGST, eds. E. Smith and A. Koratkar, ASP Conference Series v. **133**, 42.

Gilliland, R.L., et al. 2000. *Bulletin of the American Astronomical Society* **196**, 202.

Gonzalez, G., G. Wallerstein, and S.H. Saar. 1999. *Astrophysical Journal* **511**, L111.

Henry, G.W., G.W. Marcy, R.P. Butler, and S.S. Vogt. 2000. *Astrophysical Journal* **529**, L41.

Jha, S., D. Charbonneau, P.M. Garnavich, D.J. Sullivan, T. Sullivan, T.M. Brown, and J.L. Tonry. 2000. *Astrophysical Journal* **540**, L45.

Karkoschka, E. 1994. *Icarus* **111**, 174.

Kormendy, J. 2000. "Galaxy Disks and Disk Galaxies." Astro-ph/0007401.

Marcy, G.W., and R.P. Butler. 2000. *Publication of the Astronomical Society of the Pacific* **112**, 137.

Marcy, G.W., R P. Butler, and D.A. Fischer. 1999. *Bulletin of the American Astronomical Society* **194**, 1402.

Marcy, G.W., R.P. Butler, S.A. Fischer, and S.S. Vogt. 2000. IAU Symposium 202, in press.

Marcy, G.W., R.P. Butler, and S.S. Vogt. 2000. *Astrophysical Journal* **536**, L43.

Marcy, G.W., R.P. Butler, S.S. Vogt, D. Fischer, and J.J. Lissauer. 1998. *Astrophysical Journal* **505**, L147.

Mayor, M., and D. Queloz. 1995. *Nature* **378**, 355.

Merritt, D., and L. Ferrarese. 2001. To appear in "The Central kpc of Starbursts and AGN," in press. Astro-ph/0107134.

Queloz, D., A. Eggenberger, M. Mayor, C. Perrier, J.L. Benzit, D. Naef, J.P. Sivan, and S.Udry. 2000. *Astronomy and Astrophysics* **395**, L13.

Sagan, C., and E.E. Salpeter. 1976. *Astrophysical Journal Supplement* **32**, 737.

Sanders, D.B., and I.F. Mirabel. 1996. *Annual Review of Astronomy and Astrophysics* **34**, 749.

Santos, N.C., M. Mayor, D. Naef, F. Pepe, D. Queloz, S. Udry, and A. Blecha. 2000. *Astronomy and Astrophysics* **361**, 265.

Sarzi, et. al., 2000. "Super Massive Black Holes in Bulges." Astro-ph/0010240.

Silk, J., and M.J. Rees. 1998. *Astronomy and Astrophysics* **331**, L1.

Vogt, S.S., G.W. Marcy, R.P. Butler, and K. Apps. 2000. *Astrophysical Journal* **536**, 902.

Wolszczan, A., and D.A. Frail. 1992. *Nature* **335**, 145.

Chapter 3. Telescope Performance Specifications

3.1 Introduction

In this chapter we discuss several issues relating to the overall telescope performance. We present and explain the major decisions made to specify the CELT project, including the type of telescope, the optical design and the foci that will be used. We also discuss the various facilities that are needed to enable the telescope to perform well, including the image quality, adaptive optics, and the enclosure. We also include some general requirements, especially on the performance of the telescope.

3.2 Science Related Goals

From the perspective of an astronomer, any limitations placed upon the ideal will reduce the scientific potential of a telescope. Developing specifications for a telescope is an art of compromise. The major components include scientific goals, technological feasibility, start date for scientific observations, and available funds. Specifications represent a thoughtful and idiosyncratic expression of the group of astronomers who, through their astronomical interests, make difficult choices determining those specifications. There are several key issues to be considered.

First, the collecting area of the telescope tends to define the faintest object that can be studied. Most astronomical observations are photon-starved, hence the larger the collecting area the better. We select the diameter D = 30 m because we judge that this will bring major new scientific possibilities in a short time scale, and it is consistent with technical, schedule and financial realities.

A second key issue is the image quality or angular resolution that can be achieved. Many if not most astronomical observations benefit from improved angular resolution. This is fundamentally limited by diffraction to an angular resolution of λ/D. This also demonstrates a basic benefit of increasing the size of the telescope. Unfortunately, an earthbound telescope must gather light that has passed through the atmosphere, and its thermal turbulence usually blurs the image very significantly, at best to levels of about 0.5 arcsec. The importance of this atmospheric limitation drives us to search for earth-based sites that have the smallest thermal turbulence above the site, and also to the development of adaptive optics, a technique that can significantly reduce the effects of thermal turbulence on image quality.

A third key issue is the field of view (FOV) of the telescope. Often the astronomer wants to study statistical samples of groups of objects, such as galaxies; hence the ability to study multiple objects at once is a great advantage. This typically requires scientific instruments designed with this capability, as well as a telescope that can deliver high quality images over a relatively large FOV.

In many ways the most important issue influencing the power of an observatory is the capability of the scientific instruments to which the telescope delivers the light. Because we expect that scientific instrumentation for CELT will evolve over the life of the telescope, we must make our best predictions and design a telescope to support both established current needs and predicted future needs. Experience suggests that the telescope characteristics are relatively permanent.

Many other issues influence the effectiveness of the telescope: the area of sky over which the telescope can point; the speed and accuracy with which the telescope can point to an object; the transparency of the atmosphere above the telescope; the nature and magnitude of the light blockages of the telescope; the temperature and thermal emission of the telescope, etc.

3.3 General Optical Configuration

3.3.1 Basic Design Choice: Filled or Unfilled Aperture

A fundamental choice must be made of the configuration of the collecting area. Typically telescopes have a single and continuous primary mirror. However, since angular resolution increases with the baseline D of the telescope, configurations with the collecting area distributed into distinct telescopes or separated mirrors can increase D for a fixed collecting area and thus potentially achieve improved angular resolution. Such unfilled apertures are sometimes chosen. However, they increase the mechanical size of the telescope and associated costs. It is also typical that the FOV of these systems is decreased. We have chosen a single, filled aperture to best fulfill our major research goals and to limit the complexity, time to completion, and cost of the facility.

3.3.2 Basic Optical Configuration

CELT will use just one focal location, the Nasmyth. In order to provide the desired FOV, we will use a Ritchey-Chrétien optical design. The primary mirror will be hyperbolic with a focal length of 45 m. The secondary mirror will be sized to provide an unvignetted FOV of 20 arcmin at a final f-ratio of f/15. The secondary mirror position will be actively controlled, and to reduce any low frequency image motion from facility-induced vibration or windshake, the secondary should have tip-tilt closed-loop control at a bandwidth of 1 Hz.

In addition to the f/15 secondary, provision may be provided for an adaptive secondary that can deliver wavefront compensation to produce Strehl-ratio $S \geq 0.9$ at 5 μm. This will require at least 500 actuators.

For engineering reasons the telescope will be an altitude-azimuth mechanical mount. In order to maximize the size of the Nasmyth platforms and to minimize the size of the enclosure, we require that the elevation axis be in front of the primary mirror. This choice will be carefully reviewed in the next phase of work.

3.3.3 Nasmyth Foci

The Nasmyth foci offer the same image quality as the Cassegrain focus. However, these locations provide space for large instruments, and the gravity vector does not change relative to the platform. We will provide two Nasmyth platforms, each approximately 15 m x 30 m (450 m^2) in size. Each platform will have a load carrying capacity of 80 tons. A disadvantage of the Nasmyth focus is that the flat tertiary mirror, used to fold the light to the Nasmyth platforms, will absorb additional light.

The tertiary mirror will be tip-tilt articulated to provide direct access for more than two instruments. This allows the tertiary to direct the light to a range of instrument locations both on and off the elevation axis without any additional mirrors.

Polarization measurements are more difficult at the Nasmyth, so efforts will be made to minimize the degree of polarization of the light reflected from the tertiary (by maximizing the reflectivity), and to minimize its unpredictability by mapping it as a function of elevation angle.

Normal facilities will be available to support these instruments. This includes handling equipment, cooling capabilities, clean electrical power, support of cryogenic cooling systems, and guiding systems. Field rotation is a natural phenomenon at a Nasmyth focus. Compensation for this will be the responsibility of the scientific instrument, as will atmospheric dispersion compensation.

3.3.4 Prime Focus

We mention briefly two alternate foci that have been rejected. First and simplest is the prime focus, where a single reflection from the primary mirror brings the light to a focus. This focus would have the smallest light loss from reflectivity of the optics, and also have the smallest image sizes in a linear sense due to the relatively short focal length. However, this focus has several significant disadvantages. First, images at prime focus suffer from coma; thus the uncorrected FOV is small. For CELT, coma grows linearly with field angle to 0.5 arcsec at a field angle of 6.0 arcsec. Second, providing a prime focus capability requires both correcting optics and a scientific instrument, and these significantly increase the length of the telescope, thus increasing the required size and cost of the telescope enclosure. Third, servicing this focus would be inconvenient and expensive. Fourth, the primary mirror for a Ritchey-Chrétien design is hyperbolic, producing images at prime focus with about 1.3 arcsec of spherical aberration. As a result of these disadvantages and our view that the potential advantages are modest, we will not have a prime focus.

3.3.5 Cassegrain Focus

The second alternate focus is the Cassegrain focus. The Cassegrain focus requires two reflections, hence additional light loss over the prime focus. With a Ritchey-Chrétien focus, it is possible to eliminate coma as a field-angle-dependent aberration. Thus the Cassegrain focus overcomes many of the disadvantages of the prime focus. However, its location directly behind the primary mirror makes it physically awkward to access instruments. Experience at Keck, where both a Cassegrain and Nasmyth foci exist, has shown that the Cassegrain focus can be troublesome and lead to lengthy instrument exchange times. Thus, although it has certain virtues, we reject the Cassegrain in order to simplify the design, construction and operations of CELT. We think the Nasmyth will adequately replace it.

3.4 Optical Specifications

3.4.1 Collecting Area

The primary mirror will have the collecting area of a 30-m telescope. Since the primary will be composed of hexagonal segments, the array will not be perfectly circular. However, its collecting area will be 707 m^2 with minimal blockages, amounting to no more than 10%, with a goal of 5%. Diffraction-limited performance should be as good as a 30-m diameter telescope. In addition, due to its segmented primary, diffraction effects from the segment edges will have an adverse influence on the final PSF. These effects should be minimized.

3.4.2 Seeing-Limited (Adaptive Optics Off)

CELT will be used both with and without adaptive optics. In the seeing-limited regime, the atmosphere sets limits on how good the image quality can be. The seeing set by the atmosphere is quite variable, and is also dependent on the exact site (not yet selected), so setting requirements on image quality is a probabilistic issue. We have chosen the seeing conditions of Mauna Kea as the basis for our image quality specification. We want the telescope to rarely limit the image quality. We require that the telescope optical system deliver to the scientific instruments image diameters that are degraded by no more than 10% above the 10 percentile best seeing. For Mauna Kea, the average image size (500 nm wavelength, FWHM image size) is 0.5 arcsec, and the 10 percentile image size is 0.30 arcsec. From this we set a specification that the telescope (in the absence of the atmosphere) deliver to science instruments images with 0.137 arcsec FWHM. Assuming the images are atmosphere seeing-limited,

this is equivalent to a one-dimensional rms of 0.059 arcsec or 0.259 arcsec 80% enclosed energy diameter. The achieved image size will generally be wavelength dependent, but here we are making a geometrical optics approximation.

3.4.3 Adaptive Optics On

We expect that adaptive optics (AO) will be commonly used on CELT, and that the resulting images will be close to diffraction-limited. In this regime, the metric of rms wavefront error is an excellent predictor of image quality. A perfect telescope would deliver perfect diffraction-limited images, and any degradation from this is commonly measured by the Strehl ratio, S, the peak intensity divided by the peak intensity for a perfect telescope. For small wavefront errors, $S(\lambda) = \exp(-(2\pi\sigma/\lambda)^2)$ where σ is the rms wavefront error (nm) and λ is the wavelength of light (nm). At this time it is impractical to build an AO system that effectively removes the wavefront errors introduced by the atmosphere for all wavelengths. Thus our requirements for AO are set by our estimate of what capability in AO might be achievable in the next decade. This is an arena where better performance will directly benefit our science. A careful balance between our desires and what we can expect state-of-the-art engineering to deliver sets the requirements. Our AO-on specification is that the telescope deliver images with a Strehl of 0.5 at a wavelength of 1 μm. This is equivalent to a residual wavefront error of 133 nm. Since the AO performance is very sensitive to the atmospheric conditions, this requirement is set for the expected median seeing conditions, $r_0 = 0.2$ meters at a wavelength of 500 nm. This is equivalent to seeing of 0.5 arcsec. To make this requirement complete we must also specify the assumed atmospheric time constant, here set as an expected median value of $\tau_0 = 2.7$ milliseconds. In practice the adaptive optics system itself will correct or reduce many sources of telescope wavefront error, so this requirement must be imposed on the telescope AO system, not just on one of these components.

Since the achieved Strehl ratio will depend on wavelength, the performance for science will depend on wavelength in a critical way. Figure 3-1 shows the Strehl ratio as a function of wavelength. Many science programs can be effectively carried out with Strehl ratios of 0.2 or more, so one can expect the AO system to be used for observing programs with λ as short as 650 nm depending on the actual atmospheric conditions and the nature of the science program.

Achieving excellent Strehl ratios is important over the entire sky. Thus we require that this be achievable over most of the sky. At this time this requirement implies the need for artificial beacons, as the number of sufficiently bright natural stars is insufficient. We anticipate that this will be achieved with sodium beacons.

A large FOV is also desirable. We require that the AO system and telescope deliver the required Strehl ratio images over a 2-arcmin FOV at 2 μm. Since the FOV will vary as $\lambda^{6/5}$ the well corrected FOV varies with wavelength, but we will still get a large FOV at 1 μm of 52 arcsec.

In order to allow for future upgrades, we also require that the telescope be designed so as not to preclude Strehl-ratio > 0.5 at 0.5 μm with some more powerful AO system.

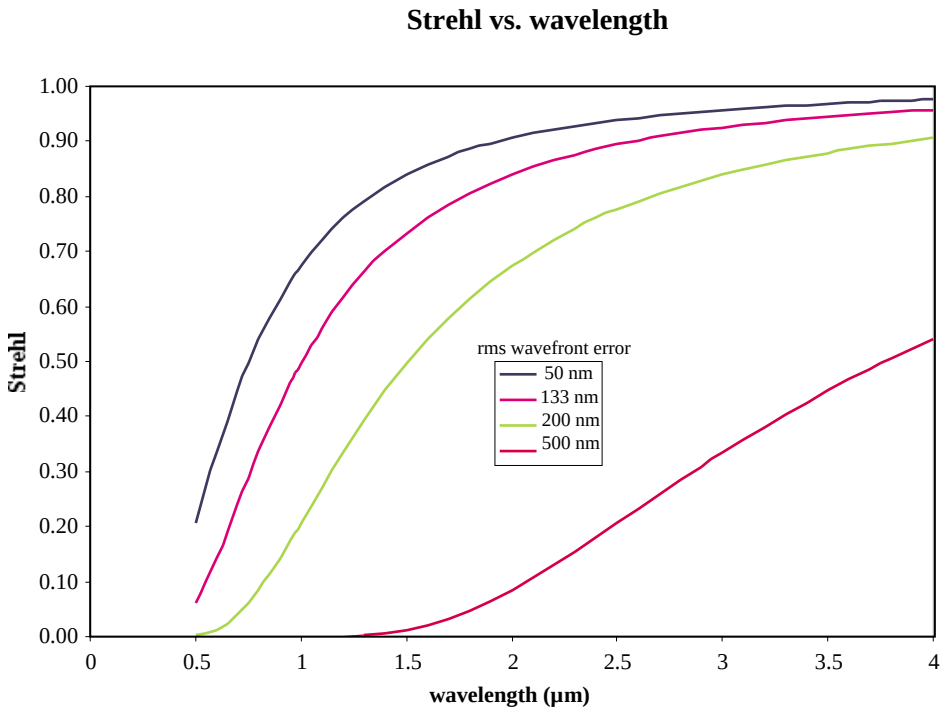


Figure 3-1. Strehl-ratio vs. wavelength for various assumed rms wavefront errors.

3.4.4 Criteria for Thermal Infrared

In general both the telescope and the atmosphere will radiate due to their thermal properties, but these fluxes are not a limitation until the wavelength is longer than about 2 µm. Beyond this wavelength, details of the telescope and instrument design can have a powerful influence on the instrument sensitivity, as limited by the thermal flux. The issues are perhaps most clearly demonstrated with observations in the 10 µm atmospheric window. Here the overwhelming background seen by any science instrument will be from thermal radiation from the atmosphere, the telescope and its optics. The atmosphere at its most transparent will have an emissivity of about 0.02. Each additional optical surface will have an emissivity of at least 0.01 depending on the cleanliness of the surface. Thus any addition of warm optical surfaces will noticeably reduce the sensitivity in the best atmospheric windows. Unfortunately, at other wavelengths the atmosphere is not as transparent, so the atmosphere background flux will be higher making the optics relatively less important. The primary mirror segment edges and gaps will also be a source of thermal emission, as will any structural blockages in front of the primary.

Thus we want several things: maximum mirror reflectivity (emissivity = 1 - reflectivity), minimal reduction of the primary mirror area by segment edges and gaps, minimal number of warm mirrors before the science instrument, and minimal blockage of the primary by the structure needed to support the secondary mirror. In practice, re-imaging the blockage onto a cold blockage inside the science instrument can largely eliminate the effect of geometrically simple blockages. Thus we want any unavoidable blockage to be as geometrically simple as is practical.

The present design has three mirrors needed before the instrument focus, and with AO this number may be significantly larger, depending on the AO design and location.

We also set a requirement that the segment edge/gap blockage be under 1%. We set as a goal that for well-designed instruments the system emissivity at 10 μm should be no more than 4%.

3.4.5 Wavelength Coverage

The atmosphere is transparent from about 300 nm to 28 μm. Ozone blocks radiation below 300 nm making the atmosphere opaque. From 10 μm to 28 μm water vapor plays an increasingly strong role in creating absorption features. Beyond 28 μm water vapor makes the atmosphere completely opaque except from extreme altitude sites or airplanes. Figures 2-3, 2-4, and 2-5 show the transmission of the atmosphere. Over this range of wavelengths where the atmosphere is transparent, we require that the telescope optics have high reflectivity, hence high throughput and low emissivity.

3.4.6 Pointing and Tracking

The telescope must accurately and quickly point to and track the scientific targets. Some programs require observing many objects per night, so speed of acquisition can be important. We have set requirements for offsetting, tracking and pointing that will make CELT an effective telescope, while not introducing prohibitive requirements.

Pointing	within 1 arcsec rms over all sky	
Tracking (open-loop)	0.02 arcsec rms in 10 seconds	
	0.1 arcsec rms in 10 minutes	
	0.5 arcsec rms in 1 hour	
Guiding (closed-loop, AO off)	within 0.02 arcsec rms over 10 minutes	
	within 0.05 arcsec rms over 1 hour	
	focus within 0.02 arcsec	
Slewing	360° azimuth, 65° elevation in 5 minutes	
	1 arcsec on sky	1 s
	10 arcsec on sky	3 s
	100 arcsec on sky	10 s
	1000 arcsec on sky	30 s
	slewing **goals** are 50% of above values	
Sky coverage	unvignetted above 25°, >75% of az range	
	< 2° zenith blind spot diameter	
Observing range		
Azimuth	100.5° ±220.5°	
Elevation	25°-89°	

3.5 Environmental

3.5.1 Solar

During the day routine scientific observing is not expected. During this time, the enclosure should shield the telescope environment against direct sunlight and insulate this environment against temperature increases.

3.5.2 Thermal

The observatory should operate within specifications over a temperature range corresponding to 90% of the time (as an example, for Mauna Kea this is $2° \pm 4°$). Normal operations should be possible over the temperature range corresponding to 99% of the time (for Mauna Kea this is $2° \pm 16°$). No damage should occur to any part of the observatory over a range $0° \pm 35°$.

3.5.3 Wind, Precipitation, and Dust

Although our site is not yet known, we pick the operational wind conditions based on Mauna Kea. For Mauna Kea, median wind speed is about 7m/s, and the wind is less than 12 m/s 90% of the time. We expect the telescope to meet its specifications in 90 percentile winds.

The observatory will experience a range of weather conditions. With the enclosure closed, the observatory should survive 100-year storms that will probably have winds up to 70 m/s. Snow and freezing rain are also expected and should not damage the observatory. Ice loads up to 0.1 m should not damage the observatory. During high winds, dust is also expected. Under these conditions, the enclosure will be closed and sufficiently well sealed that dust penetration will not be a concern. Operations should be achieved at up to 98% relative humidity. Condensation on optical and other sensitive surfaces must be avoided.

3.5.4 Seismic

The observatory should adhere to all applicable codes and survive100-year earthquakes without significant damage. The observatory will likely be built in a seismically active zone.

3.6 Operational

3.6.1 Telescope Enclosure

We require that the enclosure motion can keep up with the telescope motion so the enclosure never interferes with observing. The enclosure rotation range should be unlimited.

Achieving excellent image quality also requires good control over the local thermal environment. The enclosure should be designed to provide excellent thermal protection and control of the telescope during the day, holding the telescope at the expected nighttime temperature, and not inject any significant heat into the enclosure air at night.

3.6.2 Telescope Servicing

The observatory will be provided with all required equipment for major and routine servicing of the telescope. The construction firm shall provide equipment needed for the erection of the telescope.

3.6.3 Instrument Handling

The observatory will have suitable handling equipment and work space to carry out routine servicing of the instruments as well as installation and removal tasks. Clean laboratory space suitable for upgrades and repairs to the instruments will be provided.

3.6.4 Segment Storage

The segmented primary will be made of 1080 segments. Since each segment type occurs 6 times, we plan to make spares for each type, for a total of 180 spare segments that must be stored in the local support facility. Other mirror exchange and handling apparatus must also be supplied.

3.6.5 Mirror Covers

Although many smaller telescopes have mirror covers to protect the primary against accidental damage and dirt, the cost and complexity of a mirror cover for CELT is prohibitive and will not be required. Since the secondary mirror points down, it is much less subject to damage or dust deposition and will not require any protection. The tertiary mirror will be uplooking, and thus dust protection is desirable and practical. We plan to have a rapidly deployable and retractable mirror cover for the tertiary.

3.6.6 Mirror Coatings and Reflectivity

The highest practical mirror reflectivity is of great importance, both in order to maximize the light delivered to science instruments and to minimize the system thermal emission. The challenge of providing extremely reflective mirrors is compounded by the desire to operate the telescope over the wide wavelength range from 300 nm to 28 μm. It is our goal that the mirrors have a reflectivity as high as the better of aluminum and silver over the operating wavelength range. In addition, this coating should be durable, so it can be regularly cleaned without degradation and withstand normal environmental abuse for at least 10 years.

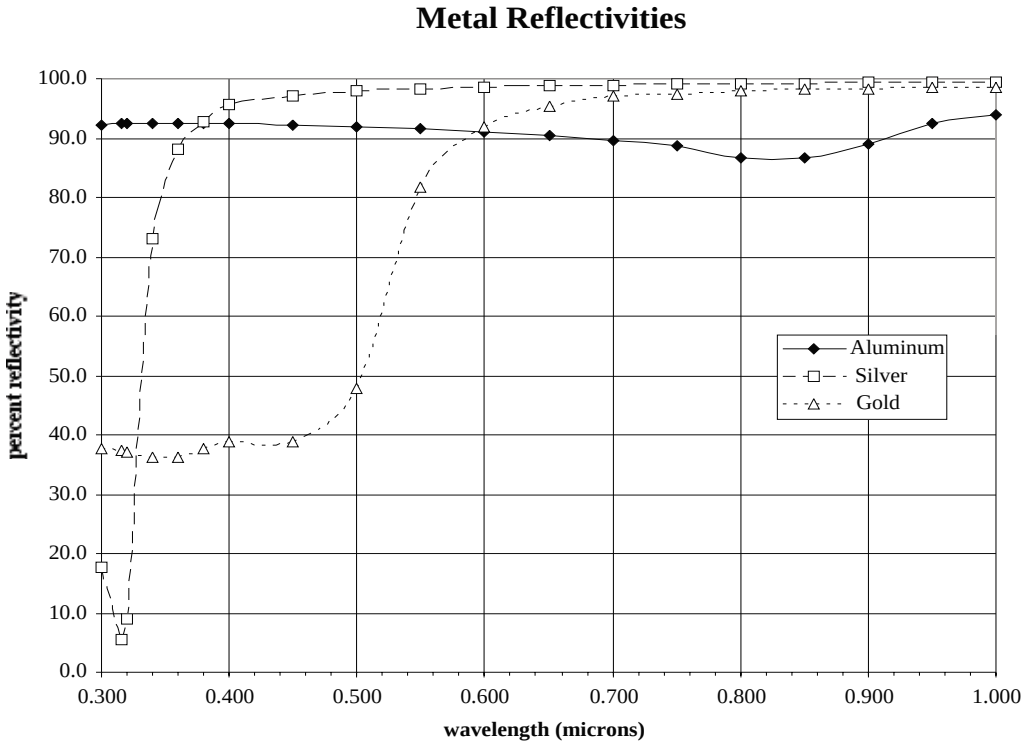


Figure 3-2. Reflectivity (percent) vs. wavelength for aluminum, silver, and gold.

Aluminum is reflective over this range, but particularly in the 300-1000 nm range the reflectivity is only about 90%. Silver is better longwards of 360 nm, but it is rapidly corroded, and it has very low reflectivity shortwards of 360 nm. Protected silver is more durable, but its reflectivity is not improved. Figure 3-2 shows the reflectivity of interesting metals over the relevant wavelength range.

Durability is particularly important for CELT. With 1080 primary mirror segments, the effort and telescope time loss due to removing all of these segments every two years is formidable. The payoff of using durable coatings with lifetimes in excess of 10 years is great.

We plan to support research efforts to develop an affordable and durable coating with extremely high reflectivity. One promising development is a high reflectivity and durable coating developed by researchers at LLNL (Thomas and Wolfe1998**).** This satisfies many of our requirements, but it absorbs strongly at 10 μm, where the atmosphere is unusually transparent. Eliminating this disadvantage will take some research effort.

3.6.7 Other Support Requirements

The facilities should provide for observing at the telescope, at a base facility, and from California. All appropriate equipment should be available for servicing of the telescope and its optics, instruments, the enclosure, other buildings and vehicles, and for the safety and comfort of personnel.

Suitable facilities for computers, networking, and backup hardware should also be provided. Support software for system administration, networking, and observatory operations should be provided. Related software and hardware for communications and remote observing should be provided.

3.6.8 Overall Reliability

It is particularly important that the entire facility be designed and built in a fashion that maximizes overall reliability and minimizes maintenance. This is important both to maximize the time available for scientific observing and to minimize operational costs. We want at least 90% of nights available for scientific observing. We have as a goal that the operations cost will be no more than 5% of the construction costs.

REFERENCES

Thomas, N. and J. Wolfe. March 2000. “UV-Shifted Durable Silver Coating for Astronomical Mirrors.” SPIE Proceedings **4003**, 312.

Chapter 4. Optical Design

4.1 Introduction

As optimal telescope specifications, the astronomer typically wants the largest possible collecting area, the highest possible angular resolution, the smallest possible backgrounds, and the largest possible observing field of view (FOV). Attempting to meet these goals is an art of compromise, based on the technical feasibility and costs of various options. Considering both technical feasibility and cost we have elected to propose a 30-meter diameter telescope. This will provide nine times the light-gathering power of the world's largest telescope, the Keck telescope. We have selected a filled aperture to achieve this increase in collecting area, a three-times increase in potential angular resolution, and a compact overall configuration.

A relatively large FOV is also desirable, one sufficient in size to encompass large clusters of distant galaxies, for example. For sufficiently scarce objects the ability to observe more than one at a time can require arbitrarily large FOVs, but for most classes of objects, FOVs of a few arcminutes are sufficient. Since we seek superb image quality, the optical design must provide low wavefront errors over this FOV. Although adaptive optics (AO) can provide diffraction-limited images, a variety of technical and instrumental limitations suggest that beyond a few arcminutes, a diffraction-limited FOV will not be practical. Thus the optical design should allow for this and need only deliver seeing-limited images over a FOV larger than this.

It is inevitable that the scientific instruments for seeing-limited images grow in size along with increased telescope size. Seeing-limited images will necessitate extremely large scientific instruments, therefore we must provide suitably large areas for scientific instruments. This makes Cassegrain foci less attractive, and encourage Nasmyth platforms that are outside of (rather than behind) the primary. Diffraction-limited images (with AO), on the other hand, do not grow with increasing telescope size, so there is the potential that instruments served by AO will be relatively compact.

The focal length of the primary mirror is a key driver for the size (and cost) of the enclosure. Shorter primary focal lengths are desirable in this regard. Unfortunately, several aspects of optical fabrication and support become more difficult with shorter primary focal lengths. In particular, the segment asphericity, a good metric for the difficulty of polishing, varies as the inverse cube of the focal length. Segment alignment tolerances also tighten as the inverse cube of the focal length. Thus the choice of primary focal length is a compromise between enclosure cost and the costs of fabricating and aligning the primary mirror segments.

The choice of final focal length also requires a compromise. Since it sets the final plate scale, the initial parts of the scientific instruments will be more compact with a shorter final focal length. However, a shorter final focal length is achieved by shortening the primary focal length and/or using a larger (more expensive) secondary mirror.

In the region of thermal-infrared wavelengths ($\lambda > 3\,\mu m$) and at those wavelengths where the atmosphere is very transparent, the thermal emission from the telescope and its optical surfaces can become the dominant background. For observations at those wavelengths we want to minimize the number of warm optical surfaces and keep those surfaces as clean and reflective as possible.

Finally, the optical system should deliver images that are compact and without large wings in the point spread function. This imposes constraints on the quality of the optical fabrication.

4.2 CELT Optical Design

A telescope with a parabolic primary can focus starlight with a single reflection from the primary mirror to the prime focus. It is possible to add a secondary hyperbolic mirror near the prime focus and produce a two-mirror Cassegrain focus. An additional flat mirror can then be used to divert this focal location to either a Nasmyth focus or a bent Cassegrain focus. The addition of more mirrors can be used to improve some optical parameters, but at additional cost, complexity, and loss of light. We will briefly describe the simplest options and explain why we have chosen a modified Nasmyth focus as our only focus type for CELT.

4.2.1 Ritchey-Chrétien Concept

By choosing an appropriate hyperbolic primary and secondary, the coma seen at the Cassegrain focus can be eliminated, at the expense of adding spherical aberration to the prime focus. The Ritchey-Chrétien design has primary and secondary hyperboloids that produce an excellent focus with only astigmatism as a low order aberration. This aberration grows quadratically with the field angle and only becomes objectionable at field angles larger than about 10 arcmin.

We briefly explore a range of optical designs that might be acceptable for CELT. We want a large FOV, up to 20 arcmin, and we want excellent image quality. We also want near diffraction-limited wavefront quality so as not to preclude the use of adaptive optics. Adaptive optics ultimately needs very small wavefront errors to achieve diffraction-limited imaging. Thus it is important that the optical design not introduce any wavefront errors that cannot be readily removed by the AO system. Large secondary mirrors are difficult and expensive to fabricate. A strongly curved telescope focal surface (short radius of curvature) and a mechanically large FOV increase the difficulty of making scientific instruments. The focal length of the primary, the size of the secondary, and the final focal length all influence the optical performance. In Table 4-1 we vary the primary focal ratio and the final focal ratio. For each combination we list the FOV with 0.5 arcsec images (100% enclosed energy diameter), the secondary mirror diameter, the focal surface radius of curvature, and the linear size of a 20-arcmin FOV.

Table 4-1. Possible Optical Design f-ratios

Primary f-ratio	Final f-ratio	FOV (arcmin)	Secondary Diameter (meters)	Focal Surf Radius (meters)	20 arcmin FOV (meters)
1.75	17.5	22.39	3.85	6.46	3.05
1.75	15	22.29	4.38	7.45	2.62
1.75	12.5	22.16	5.10	8.81	2.18
1.75	10	21.97	6.12	10.77	1.75
1.5	17.5	21.21	3.47	5.00	3.05
1.5	15	21.13	3.96	5.78	2.62
1.5	12.5	21.03	4.62	6.85	2.18
1.5	10	20.87	5.56	8.42	1.75
1.25	17.5	19.95	3.08	3.71	3.05
1.25	15	19.89	3.52	4.30	2.62
1.25	12.5	19.80	4.12	5.11	2.18
1.25	10	19.67	4.98	6.30	1.75
1.0	17.5	18.58	2.67	2.59	3.05
1.0	15	18.53	3.06	3.01	2.62
1.0	12.5	18.47	3.60	3.59	2.18
1.0	10	18.37	4.38	4.45	1.75

From Table 4-1 we learn several things. First, the FOV with good images is only slightly dependent on the optical parameters selected. Second, shorter primary focal lengths lead to smaller secondary mirrors. Third, the curvature of the focal surface can be large. A slower final f-ratio and faster primary f-ratio make the focal surface curvature greater. Large curvatures will make instruments that use the full field more difficult to design and build. The specific instruments discussed for CELT in Chapter 10 do not appear to be particularly sensitive to focal surface curvature.

We selected a reference design on the basis of the length of telescope, the difficulty of fabricating the primary, and the numbers in this table. The reference design has a primary f-ratio of 1.5 and a final f-ratio of 15. The major parameters that flow from this design are listed in Table 4-2.

Table 4-2. Reference optical design assuming an f/1.5 primary and f/15 RC design

Parameter	**Value**
Primary mirror diameter	30 m
Primary mirror focal length	45 m
Final focal ratio	15
Final focal length	450 m
Back focal distance (behind primary)	16.5 m
Conic constant of primary	-1.002837
Conic constant of secondary	-1.525154
Radius of curvature of secondary (convex)	-12.42424 m
Primary-secondary distance	39.40909 m
Secondary mirror diameter for 20 arcmin FOV	3.95655 m
Secondary mirror diameter for 0 arcmin FOV	3.72727 m
Final plate scale	2.18 mm/arcsec
Size of 20 arcmin FOV	2.618 m
Radius of curvature of focal surface	5.78 m
Tertiary mirror distance in front of primary	3.5 m
Diameter of a single star’s footprint	1.333 m
Minor diameter of tertiary	3.097 m
Major diameter of tertiary	4.380 m

It is important to note that these parameters have been specified for a 30-m circular primary. The segmented nature of the primary will cause it to be non-circular. This leads to ambiguity about the meaning of some of these parameters. In certain cases, the circumscribing diameter will be more appropriate as input than the effective diameter. The current segmentation design has a circumscribing diameter of 31.0 m. Figure 4-1 shows the selected optical configuration.

The Ritchey-Chrétien design gives perfect on-axis images to lowest order. To higher order the pair of hyperboloids does not give perfect correction. In Figure 4-2 we plot the wavefront error (WFE) of the reference Ritchey-Chrétien system as a function of field angle. For small angles (<1 arcmin) the WFE is sufficiently small to not adversely affect the AO system. For large angles the WFE grows in agreement with the predicted astigmatism.

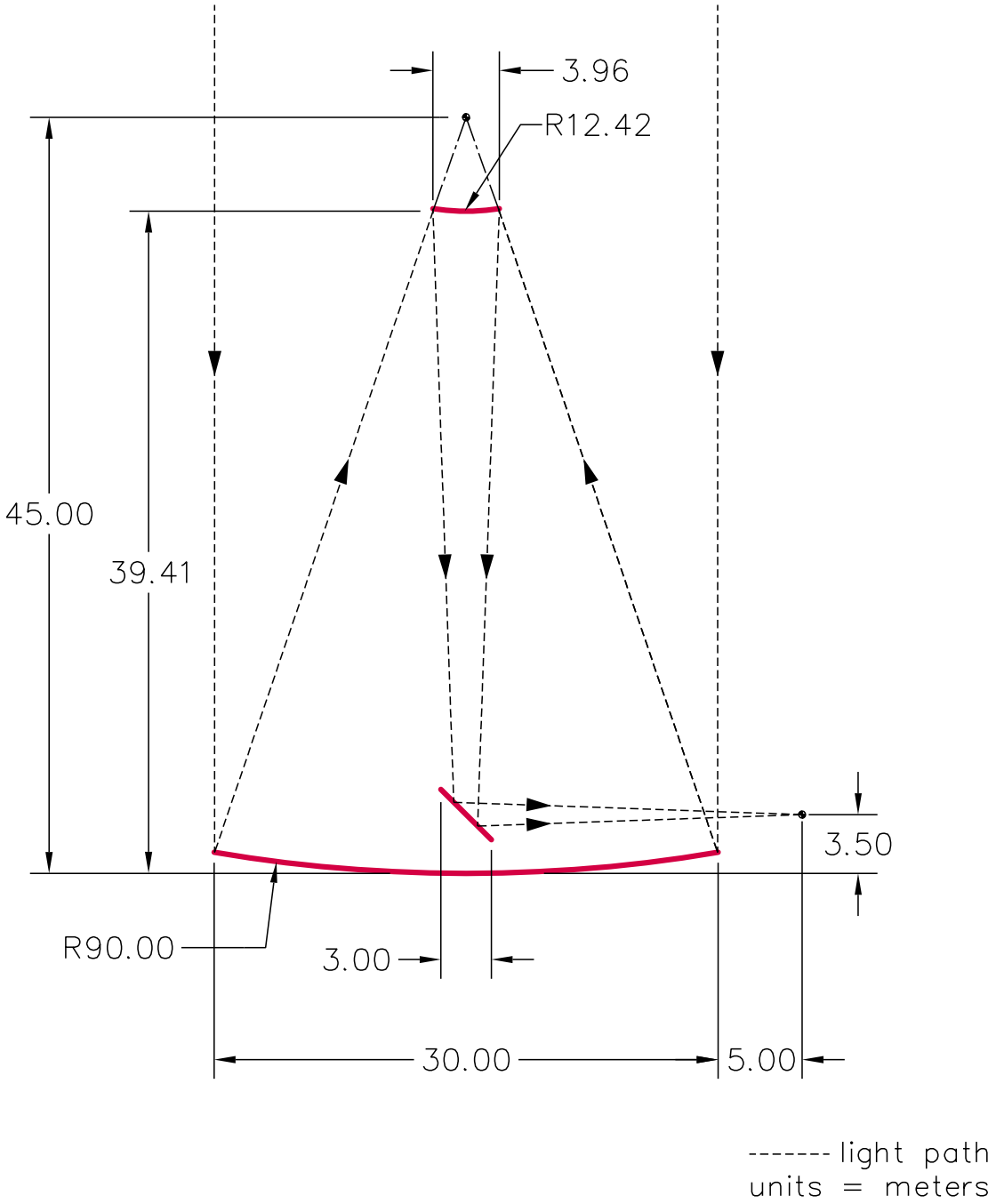


Figure 4-1. Optical layout of CELT showing the primary, secondary and tertiary, and the Nasmyth focus.

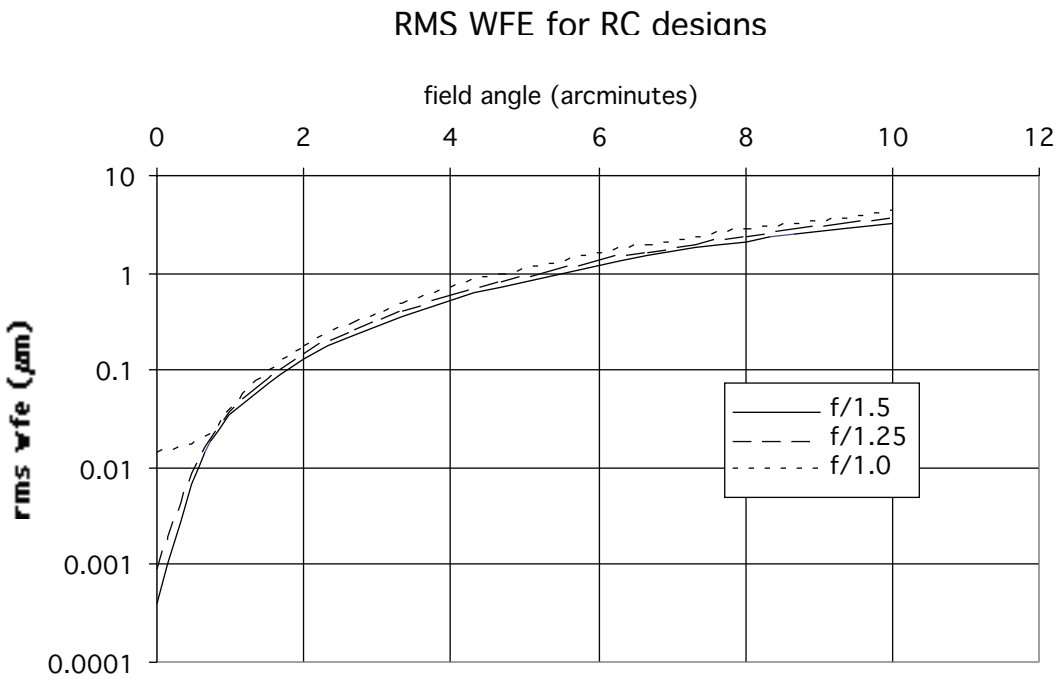


Figure 4-2. RMS wavefront error (μm) vs. field angle for the Ritchey-Chrétien design. For small angles the error is small, and for large angles the error is that expected from geometric optics.

We also show in Figure 4-3 spot diagrams of the image size for various field angles

4.2.2 Prime Focus

Using the prime focus significantly lengthens the telescope, increases the dome size, and complicates the support facilities for accessing the instruments there. At this stage of the CELT design, we expect that our scientific needs can be met at the Nasmyth focus; hence we currently do not plan to have a prime focus capability.

With a parabolic primary, perfect images will be made on-axis at the prime focus. However, off-axis images suffer from coma that grows linearly with field angle. This severely limits the useful FOV at prime focus. The comatic image size reaches 1 arcsec at 12 arcsec off axis.

For a Ritchey-Chrétien optical design (see previous section) the primary mirror is hyperbolic, so in addition to field coma, prime focus images suffer spherical aberration. For an f/1.5 primary this produces 1.3 arcsec of spherical aberration everywhere in the field. This is not acceptable. However, because the primary mirror is segmented, it is possible to reduce this by adjusting the overall primary shape by applying suitable pistons and tilts to the individual segments. These adjustments are well within the expected range of the segment actuators, and they reduce the image size to roughly 0.3 arcsec. The outermost segments have residual focus and astigmatism of about 50 nm surface in each Zernike coefficient. These uncorrectable surface errors dominate the geometric image size

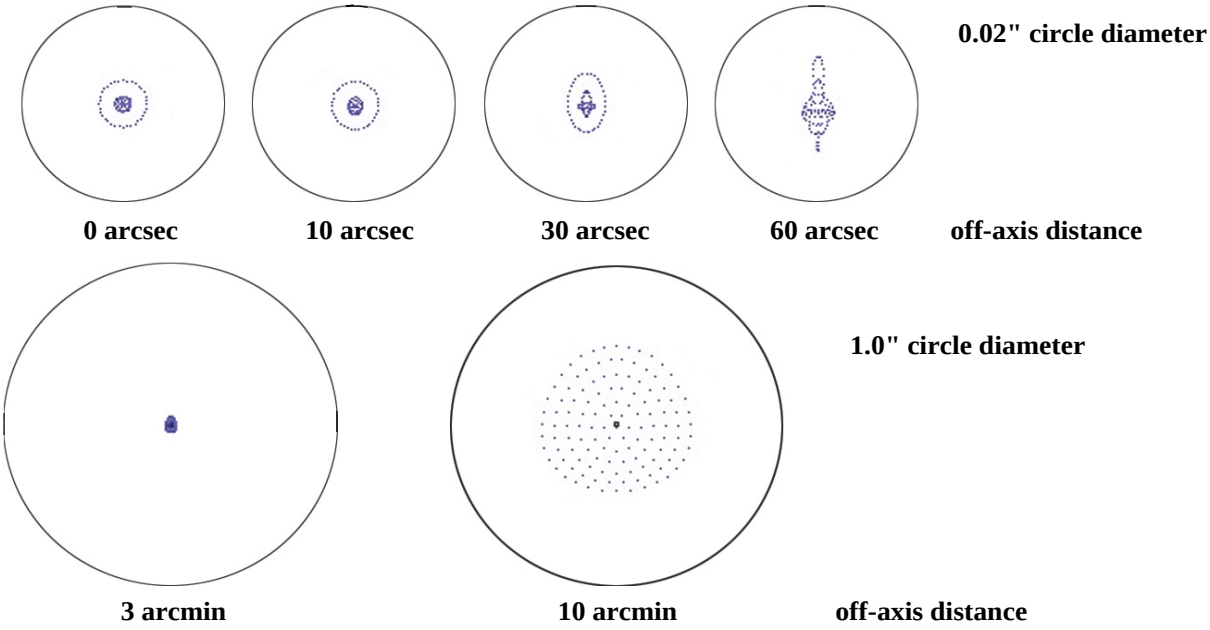


Figure 4-3. Spot diagrams for the Ritchey-Chrétien focus for field angles of 0, 10, 30, and 60 arcsec and 1, 3, and 10 arcmin.

4.2.3 Cassegrain Focus

The Cassegrain focus is created using a primary and secondary mirror. This focus is typically located behind the primary. The advantages of this focus are that it only has two reflections and that the geometry is symmetric, thus minimizing polarization. However, this focus also has disadvantages. It complicates the design of the mirror cell. The space behind the primary mirror is restricted in size, thus limiting the size of potential scientific instruments. This also adds to the complexity of servicing instruments since they are in a restricted space. The instruments must be designed to tolerate a changing gravitational load as the telescope moves from zenith to horizon. At the Keck Observatory the instruments at the Cassegrain focus must be rolled out of position for servicing or replacement. At the Gemini Observatory multiple instruments are simultaneously available at the Cassegrain focus, but their size is even more restricted. Given these disadvantages, we have rejected the Cassegrain focus for CELT.

4.2.4 Nasmyth Focus and Platforms

A third flat mirror can be added to reflect the Cassegrain focus to another location, typically a point on the elevation axis. This location is called the Nasmyth focus. At this location, it is often practical to add a large platform to hold scientific instruments. When the focus is on the elevation axis of the telescope, the focus is stationary as the telescope tracks a star. However, with altitude-azimuth telescope motion the field rotates, as it does at Cassegrain.

Two advantages of the Nasmyth focus are that there is usually ample space available for scientific instruments, and the image on the elevation axis does not move as the telescope tracks a star. In addition, the gravity orientation on the Nasmyth platform is fixed. Its major disadvantage is that a tertiary mirror is needed, and it causes additional light loss and increased polarization effects. Because of the light loss, it is particularly important that CELT have very high reflectivity optics.

A key element needed to fully define the optical and mechanical configuration is the location of the tertiary mirror. Obvious choices include placing the tertiary mirror in front of the primary (as is done on Keck) or placing it behind the primary mirror (done on some radio telescopes).

Assuming the tertiary mirror lies on the intersection of the optical axis and the elevation axis, placing the tertiary mirror behind the primary has some optical consequences. The system must provide adequate space for Nasmyth platforms that can accommodate the desired instruments. As a result, placing the tertiary immediately behind the primary requires that the support structure have an opening for the optical beam to allow the beam to get to the edge of the primary. An alternate is to place the tertiary well behind the primary, where there might be sufficient room for the Nasmyth platforms.

It is important to note that the tertiary mirror location defines the elevation axis of the telescope. Thus the length of the telescope depends on this location. Solutions with the tertiary mirror behind the primary will generally require larger enclosures for the telescope, a disadvantage. If the primary mirror focal length is sufficiently short this may not be an issue. We will explore this possibility in the next design phase.

Placing the tertiary mirror in front of the primary allows for a more compact enclosure, and readily allows the beam to carry to the edge of the primary where the Nasmyth platforms can be located. Furthermore, at this position an articulated tertiary is possible, allowing the tertiary mirror to direct the light to several instruments on the Nasmyth platform

For this reference design, we assume the tertiary is in front of the primary, for the reasons outlined above.

Articulated tertiary to move focus

In order to make maximal use of the potential space on the two Nasmyth platforms, we propose to actively control the tertiary tip and tilt. This allows us to reflect the Cassegrain focus onto a wide range of locations on the Nasmyth platform. In principle the locus of foci forms a spherical shell centered on the tertiary mirror. Except for the focus on the elevation axis, maintaining the focus at a fixed spatial location will require continuous control over the orientation of the tertiary mirror. Analysis of this system shows the motion requirements for the tertiary are modest and do not add any unexpected complications (Kuhlen 2001). Figure 4-4 shows the range of foci available at the Nasmyth platforms with this configuration. It is also possible to move the focus vertically as well as horizontally, so vertical layers of instruments at the Nasmyth platform become practical, all with only three reflections.

Motion of secondary to shift focus along optical axis

The focus on the Nasmyth platform can also be moved to some extent along the optical axis. If the secondary is pistoned by δd, the point of best focus moves by $\delta e = \delta d(1+m^2)$, where the magnification, $m = 10$ for the reference design. Thus the image moves 101 times the secondary motion. Unfortunately, the image also acquires spherical aberration as it moves away from the design focus (which has zero spherical aberration). This aberration can be calculated with the standard ASA formula (Faber 1981; Schroeder 2000) by noting that both the magnification and back focal distance change with secondary position changes. For the CELT reference design, spherical aberration grows at a rate of 0.425 arcsec/m away from best focus.

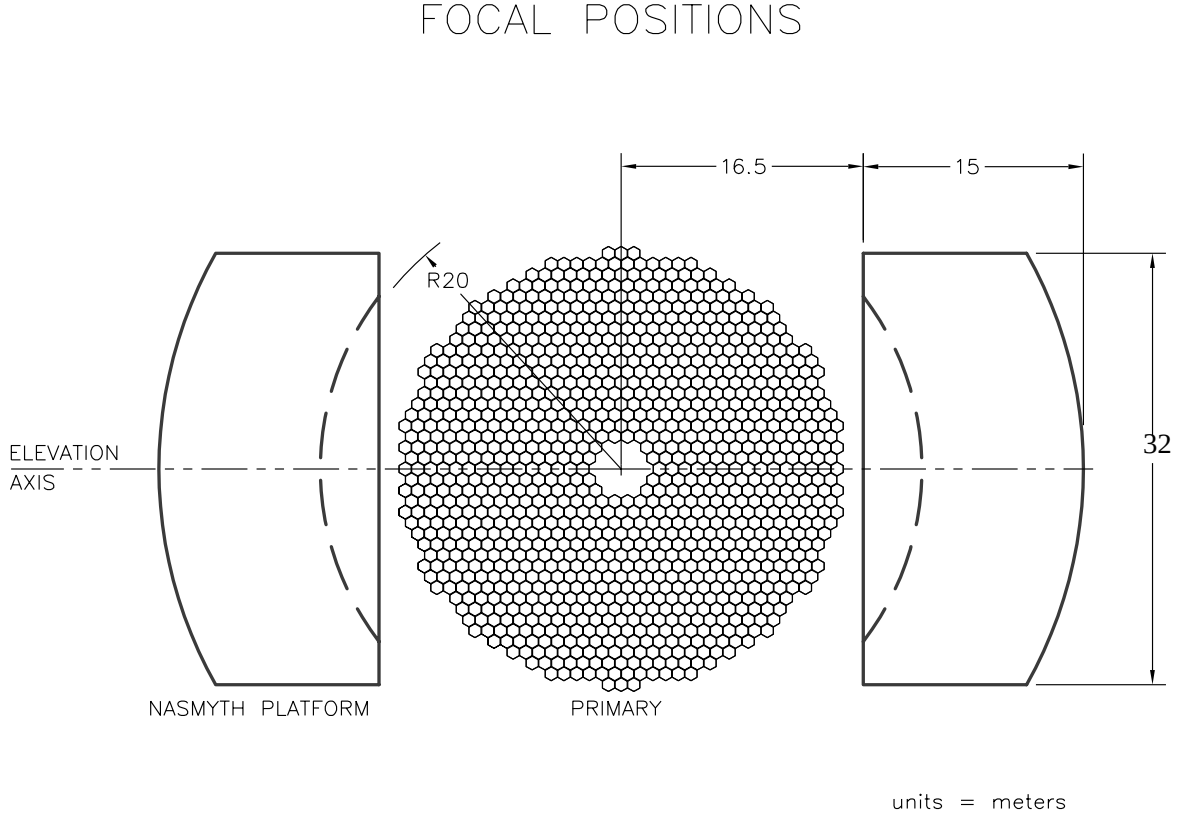


Figure 4-4. The Nasmyth platforms are shown along with the primary mirror. Also indicated as a dashed curved line is the range of focus locations possible with an articulated tertiary mirror.

Adjustment of primary segments to shift focus along optical axis

The dominant image aberration caused by moving the secondary is spherical aberration, which occurs over the entire FOV. However, the primary mirror for CELT is segmented and actively controlled. Suitable piston and tilt of the individual segments can adjust its overall shape. By tilting the segments to superpose their geometric images, one can reduce "spherical aberration" to the amount in the most-affected segment. By calculating the rate of change of slope errors across a segment and picking the outermost segment, one can determine the residual image diameter to be $3\sqrt{3}(a/R)$ ASA, where a is the segment radius and R is the primary mirror radius. For the reference design this reduces the image diameter to about 17% of ASA, or 0.074 arcsec/m. A focal shift of 3 m gives a 0.22 arcsec-diameter image.

Chopping and nodding of secondary

For various reasons, excellent sky subtraction in the infrared often benefits from image chopping or nodding. In other words, moving the star field back and forth small amounts on the detector is useful to accurately subtract the sky or other uniform backgrounds. This is typically done at a relatively high frequency, in order to eliminate low frequency drifts of detector sensitivity. For CELT we expect that we will want to move the images up to 10 arcsec and at frequencies of up to 1Hz. Since the purpose of the image motion is to aid in background subtraction, it is best if the starlight of interest not move on the optics. Thus, tilting the primary is the best approach. This is rather difficult, so the next best option

is to tilt the secondary. Chapter 6 discusses the requirements this imposes on the secondary. A third option of tilting the tertiary may also be considered, although this produces the greatest change of optical footprint on the secondary and primary, hence is most subject to non-common-path errors.

4.3 Diffraction

4.3.1 Circular Primary

The image quality of the telescope is limited by the atmosphere, the quality of the optical figuring, the optical design, and most fundamentally, by diffraction. For a circular aperture the angular size of the diffraction-limited image is given approximately by λ/D where λ is the wavelength and D is the diameter of the mirror. Thus the image size is limited to be no smaller than this. The simplest aperture is a circular one and the diffraction-limited image from this is called the Airy pattern, an axisymmetric pattern shown in Figure 4-5 for a 30-m diameter and a wavelength of 1 μm. The figure also shows the diffraction pattern of a 1 m circular mirror, approximating the diffraction pattern of a single segment. Changing the details of the aperture shape or blockage of the aperture influences the diffraction pattern, usually for the worse, causing the image to become less centrally concentrated. This spreading of the image reduces the point source sensitivity and also reduces the capacity for detecting faint objects near bright ones.

4.3.2 Perfect Segmented Primary

We next consider the consequences of segmenting the primary mirror. We assume the segment surfaces and segment alignment are perfect and that there are no gaps between segments.

Using a mosaic of hexagonal segments for the primary causes the perimeter to be non-circular. Figure 4-6 shows the PSF for a perfect primary composed of 1080 hexagonal segments (see Figure 5-1). This primary has a central hole and a serrated edge. The image here is not axisymmetric, but cuts in the PSF along the x and y axes show the essential range of distribution. By comparing Figures 4-5 and 4-6 one can see that the effect of segmentation is small in the center of the image, and makes modest changes at larger angles

4.3.3 Effect of Including Segment Gaps

An additional complication occurs when the segments have gaps between them. This is essential for both optical fabrication (safety bevels on edges) and for assembly. As detailed in Chapter 5, we expect our segments to have 1 mm bevels and a 2 mm air gap between adjacent segments. These 4 mm non-optical regions will cause a loss of light, from 701.5 m^2 to about 695.0 m^2, or about 0.9%. In addition diffraction will redirect the same amount of light into angular scales of λ/w, where w is the loss width of 4 mm, or roughly 50 arcsec. The angular pattern produced by the individual segments will also be amplified as there are more edges. This can be seen as a series of peaks at the intensity level of about $1x10^{-5}$. The repetition period is caused by interference between sets of edges, with a spacing of $\sqrt{3}a/2$. The diffraction pattern including gaps is shown in Figure 4-7. Comparing with Figure 4-6 we see that in the image core there is very little difference, but on large angular scales, the wings are significantly brighter than without gaps. Both the angular scale and the brightness of these wings are directly dependent on the gap size and gap series spacing.

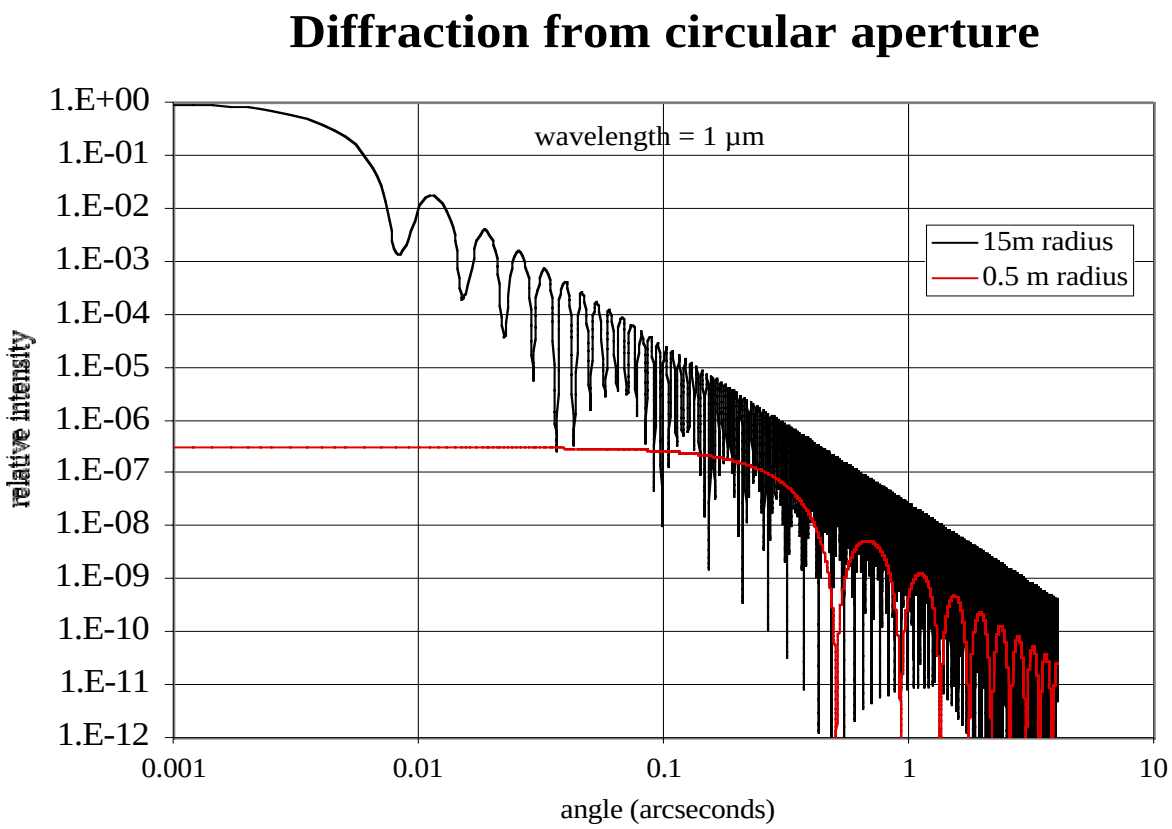


Figure 4-5. The intensity as a function of angle for a 30 m circular aperture is shown. The assumed wavelength is 1µm. Also shown is the wider diffraction pattern for a 1 m aperture, indicative of a segment.

4.3.4 Effect of Secondary Support Structure

There will be structural elements supporting the secondary mirror that will block part of the primary. The reference structure is described in Chapter 7, and the expected blockage is about 19.9 m^2 or 2.8%. Light will be lost at this level and an equal amount diffracted into angular scales defined by the sizes of the obstructions. Figure 4-8 shows the PSF due to these assumed obstructions. For the calculations we assumed three radial members at 120° clocking angles that are 0.40 m wide, and six almost radial arms that are 0.038 m wide. The geometry is shown in Figure 7-5.

4.3.5 Overall Effect

In all we expect about 26.4 m^2 of the primary will be obscured. We expect the degradation of image quality due to this will be virtually invisible except at the very outer wings of the PSF.

In the thermal-infrared, the blockages will appear as black objects, radiating as black bodies at the ambient temperature. This additional source of thermal background is potentially important, particularly at wavelengths where the atmosphere is almost transparent. A cold pupil stop in the science instrument should eliminate these thermal background sources. However, we assume that due to the complexity of the pattern of inter-segment gaps this source of background flux will not be masked, hence a thermal background source equivalent to 0.9% emissivity will be unavoidable.

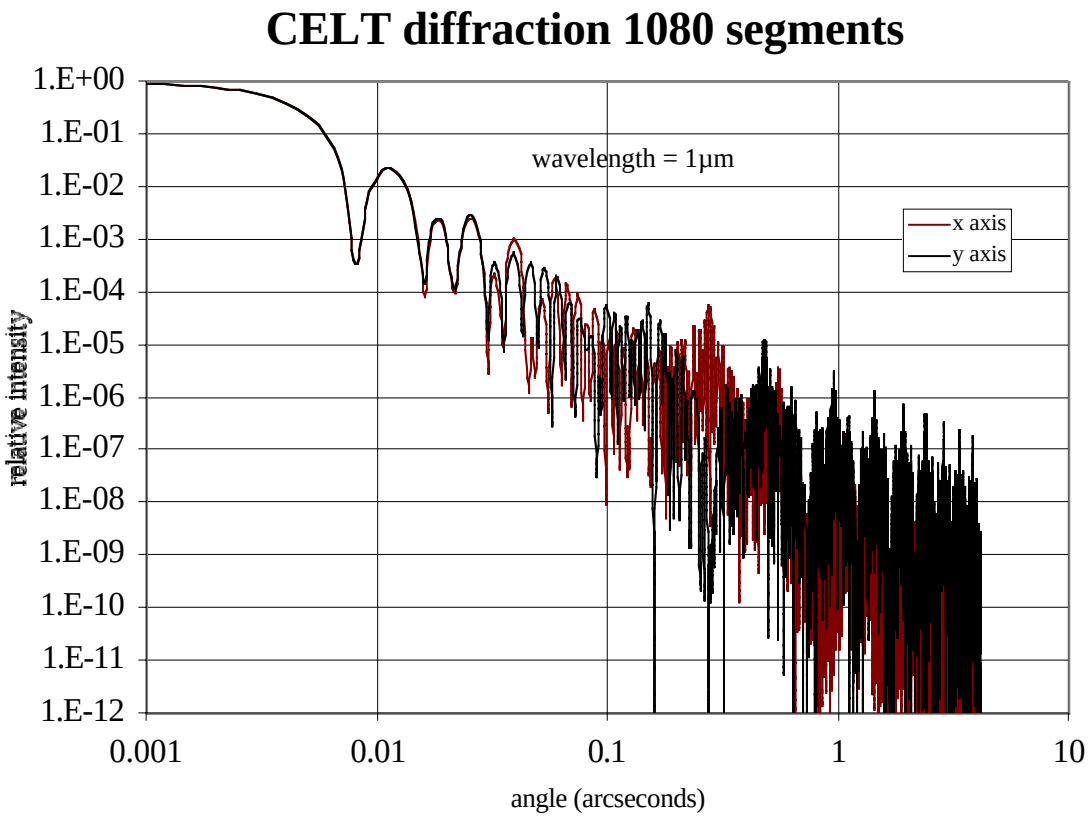


Figure 4-6. The point-spread function for the segmented primary consisting of 1080 perfect segments is shown for a wavelength of 1 µm. Cuts in the x and y axes are shown, since the pattern is no longer axisymmetric. Segment gaps have zero width in this calculation.

4.4 Effects of Misalignment

The optical design provides excellent image quality. However, misalignment of the primary, secondary and tertiary cause image motion, image degradation, and pupil motion. The position of each optic is characterized by six degrees of freedom with potential misalignments in the same degrees of freedom. In this section we address the effects of rigid body misalignments and ignore deformations of the optics.

4.4.1 Primary Mirror Misalignment

In this section we assume that the primary can be treated as a rigid body and ignore all the relative misalignments of the segments and any segment shape changes. Segment misalignments are discussed in some detail in Chapter 5.

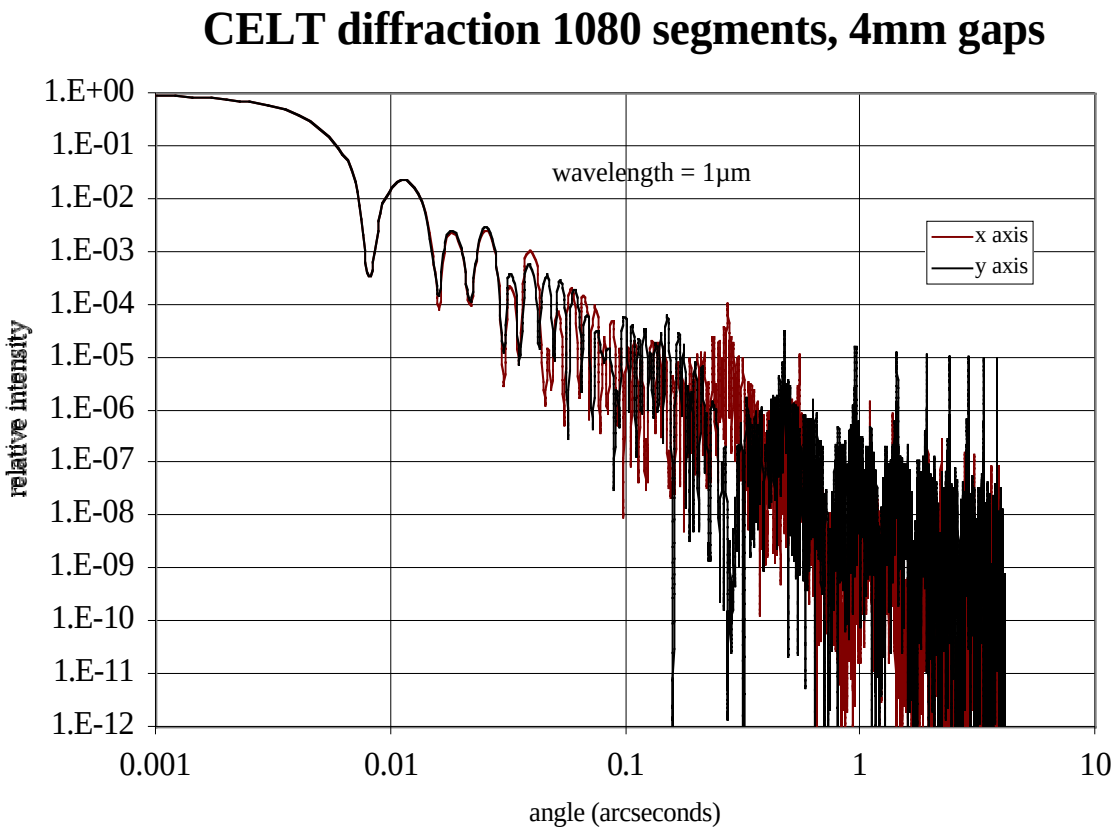


Figure 4-7. The point-spread function from the1080-segment primary with 4 mm gaps between the segments. The wavelength is 1 μm. Comparison with Figure 4-6 is interesting, showing new peaks caused by the width and interval between segment gaps. Cuts in the PSF in the x and y axes are shown.

It is important to define a coordinate system in which to define misalignment. One might think that the primary mirror itself is a natural system for this, thus by definition it is not subject to misalignment. However, since the scientific instruments and the other optics are attached to a telescope structure, not to the primary, some structural reference is also logical. Since the structure is likely to deform with gravity and thermal changes, its coordinate system is not completely well defined. However, we will discuss the primary misalignment with respect to the elevation axis defined by the yoke. This axis is likely to be fairly stable as a function of elevation angle.

The primary mirror can displace in three directions and rotate in three axes. Displacement along the optical axis is likely to be at the level of a few mm, driven by gravity. Comparable motions due to temperature effects are likely. Motions perpendicular to the optical axis will also occur, driven by gravity and temperature. The gravity effect will be of order 10 mm. The consequences of such motion, given that the other optics (secondary and tertiary) follow along rigidly, will be seen primarily as image motion at the scientific instruments. Another way of visualizing this is to recognize that this motion is identical to the instruments moving relative to the telescope. This will be automatically corrected by the guiding system. Similarly, rotations of the primary will either be benign (rotations about the optical

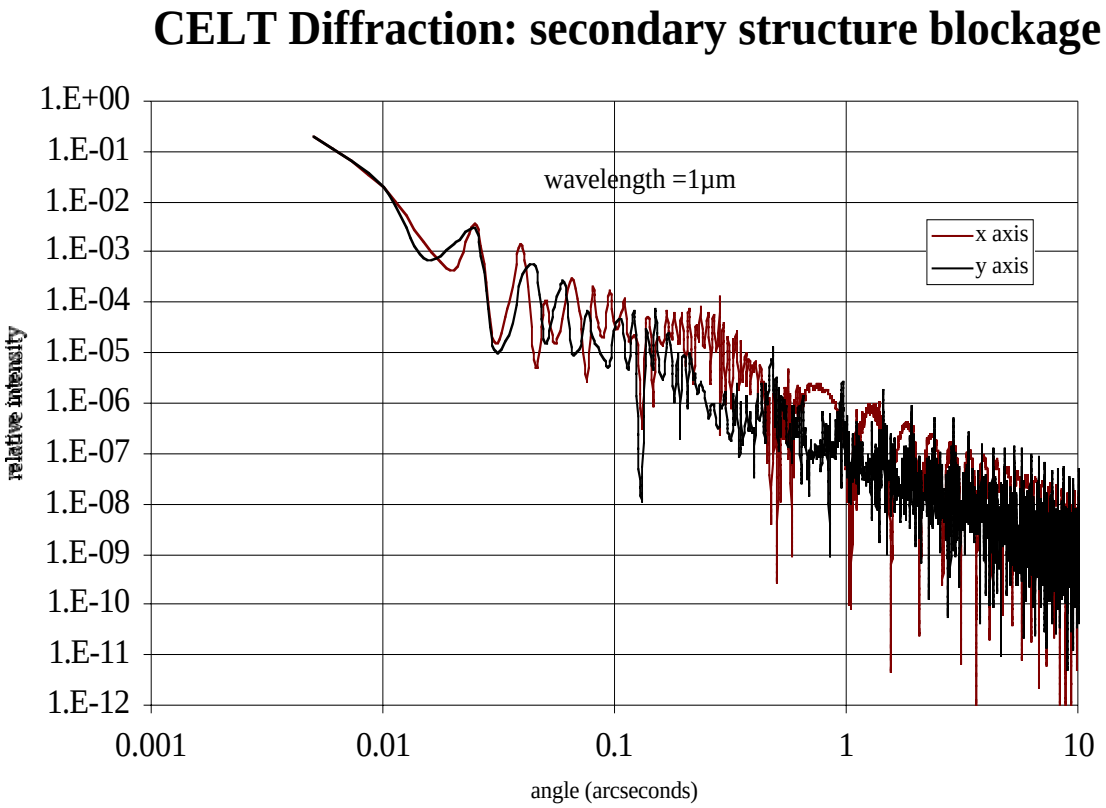


Figure 4-8. The point-spread function from the 1080-segment primary with no gaps but with the secondary support structure blockage included. The wavelength is 1 µm. Comparison with Figure 4-6 is interesting. Cuts along the x and y axes are shown.

axis), or again, cause pointing errors that are corrected by the guiding system. The pupil image is formed about 5 m behind the secondary, and is 4 m in diameter. The pupil will move the same order of magnitude as the primary itself, but since this is a small fraction of the size of the pupil, we do not expect this will be a problem.

4.4.2 Secondary Mirror Misalignment

Secondary mirror misalignment also consists of three displacements and three rotations. We define these to be relative to the coordinate system of the primary mirror.

Translations along the optical axis will shift the point of best focus, and introduce focus image blur at the nominal focus. A displacement of 1 mm will induce an image blur of 3.09 arcsec. Translations perpendicular to the optical axis will cause image motion of 4.13 arcsec/mm and misalignment coma with ATC = 0.383 arcsec/mm (Faber 1981). The tilt of the secondary will cause image motion of 0.248 arcsec/arcsec and misalignment coma with ATC = 0.010 arcsec/arcsec. A 1 mm error at the edge of the 4 m secondary produces a tilt of 52 arcsec or tilt induced coma of ATC = 0.52 arcsec. Chapter 11 suggests that ~0.04 arcsec might be allowable here. Clearly the secondary must be actively positioned to an accuracy of order 0.05 mm. We expect the gravity-driven misalignment of the secondary will be of order 20 mm.

Table 4-3. Effect of secondary misalignment

Secondary motion	Image motion (arcsec)	Image diameter (arcsec)
1 mm axial	0.0	3.09 (defocus)
1 mm lateral	4.13	0.383 (coma)
1 arcsec tilt	0.248	0.0103 (coma)

We expect that open-loop corrections of the secondary position can be made quite accurately. In addition the guider system will measure low order aberrations, such as coma, and control these misalignments using closed-loop control of the secondary position. We note here that since both decenter and tilt of the secondary produce coma, measurements of coma do not determine which combination of motions is responsible. Structural modeling will guide us here, as well as potentially an additional off-axis guider. We expect that after open-loop corrections (lookup tables) have been applied, residual corrections will be small. Thus, although multiple motions might correct the residual error, it should be sufficient to apply the corrections with the most convenient motion.

In addition, secondary misalignment will produce some effect on the plate scale and field distortion. We have not yet calculated these.

4.4.3 Tertiary Mirror Misalignment

Since the tertiary is a flat mirror, the three motions in the plane of the mirror cause no optical effects. Piston normal to the surface will produce both a shift of the focus normal to the optical axis and a shift along the optical axis. Tilt errors of the tertiary will also move the image, tilt the focal plane, and move the pupil image. We assume the x-axis is along the elevation axis and the z-axis is along the primary optical axis. Recall that 1 mm displacement at the edge of the tertiary corresponds to a tilt of about 69 arcsec.

Table 4-4. Effect of tertiary misalignment

Tertiary motion	Image motion (arcsec)	Image blur (arcsec)	Focal plane tilt (arcsec)
1mm normal piston	0.458	0.061 (defocus)	0.0
1 arcsec about y	0.089	0.0	2.0
1 arcsec about z	0.044	0.0	1.0

Because the dominant effect of tertiary misalignment is image motion, the guider will remove small alignment errors. We do not anticipate that optical-based closed-loop control of the tertiary will be required.

4.5 Additional Configurations

In addition to the Ritchey-Chrétien Nasmyth focus, we are also considering a smaller secondary for adaptive optics. The potential value of such a secondary is that it might be adaptive in its own surface, thus greatly simplifying the downstream needs of an adaptive optics system. An adaptive secondary is probably not appropriate for a high order correction system, but might be useful for AO in the thermal-

IR region, where corrections of a few hundred degrees of freedom would be very useful. Adaptive secondary mirrors are quite challenging, but at least one 0.64 m in diameter has been made for the MMT that has 336 actuators. Mirrors of this type up to 2 m in diameter might be possible in the future.

4.6 Optical Baffling for f/15 System

It is often useful to baffle optical systems. This is done to prevent stray light that is not reflecting through the optical system from getting to the focal plane. An example of the source of stray light includes the moon (or a bright star that is nearly behind the secondary mirror). Its light misses the secondary and directly illuminates the focal plane and the final detectors. This can significantly increase the background flux at the detector and degrade sensitivity.

The traditional approach to blocking this source of "bright sky" is to place a baffle around the secondary mirror and a conical baffle in the interior of the primary mirror. These can successfully baffle a Cassegrain instrument from seeing anything except light that is reflected from the primary and secondary mirror over the design FOV. The penalty is usually a modest amount of vignetting of the optics.

However, when one observes from the Nasmyth platform, the focal plane and the detectors can also see a bright dome wall and possibly parts of a bright telescope structure. The Cassegrain baffling solution would impose severe blockage of the primary mirror.

For CELT we propose that rather than impose large and incomplete baffles on the telescope itself, the scientific instruments should provide their own baffling by providing an internal image of the pupil and block all light that misses this pupil image. This assumes that scatter from intermediate optical surfaces will be sufficiently small.

We note that if the optical configuration changes to having the elevation axis behind the primary, rather than in front of the primary, it might be practical to baffle the telescope in the traditional way. Even in this case, baffles around the secondary mirror are likely to significantly increase the wind cross section of the top end of the telescope, and this could become a problem on windy nights. This configuration will be studied in the next phase of work.

4.7 Field Curvature

As mentioned earlier, the Ritchey-Chrétien focus gives a curved focal surface, rather than a flat one. For the reference CELT design the sag of the focal surface for the 20-arcmin FOV is 0.15 m. Scientific instruments using this large field must account for this curvature. The radius of curvature is indicated in Table 4-1 and was calculated using standard formulae such as those in Schroeder (2000). The curvature changes with various optical parameters. For the range of interest to us, curvature of field is very important to the instrument designer, but is not an insoluble problem.

4.8 Distortion

For two-mirror telescope systems there will also be distortion in the focal surface. This means that the plate scale varies as a function of distance from the center of the field. The error in star position (relative to the paraxial focus plate scale) grows as the cube of the field angle. The formulae for determining this can be found in Schroeder (2000), Table 6.9. For the CELT reference design at 10 arcmin off axis, the positional error is only 0.084 arcsec.

4.9 Mirror Coatings

CELT will be an expensive telescope, largely because of its great collecting area. Clearly we do not want to give up photons so expensively gathered. It is important to coat the three mirrors with the most reflective coatings available; in addition, re-coating will be difficult and time consuming. Hence, we want a coating that can be easily cleaned in place and needs to be re-coated onto the optics very rarely.

Currently at the Keck Observatory, the optics are coated with bare aluminum, and even when fresh the reflectivity per surface is not perfect (see Figure 3-2). With three reflections, no more than 70% of the visible light hitting the primary gets to the Nasmyth focus. In addition, the mirrors are re-coated every two years because even with a monthly CO_2 cleaning their reflectivity degrades with time.

For CELT we hope to use a more durable and reflective coating on all the optics. Studies by Thomas and Wolfe (2000) uncovered coatings that roughly provide the best of aluminum and silver reflectivities and also have extremely durable surfaces that should last at least a decade before they need to be replaced. It is our intention and hope that coatings of this nature will be used for CELT. These coatings have reflectivities for visible light of over 95%, thus allowing more than 85% of the light that hits the primary to be delivered to the Nasmyth focus.

It is also the case that these highly reflective coatings will be excellent in the infrared, with reflectivities beyond 1 μm of 99% or better (similar to silver).

4.10 Thermal-IR Performance

Observations beyond about 2 μm wavelength typically experience additional background light that is thermally emitted from the telescope itself, not just from the night sky or from diffuse astronomical backgrounds. Any sources that the detector can see are potential sources of this radiation, so all aspects of the telescope and instrument optics and supports become important. For example, at 0°C, objects radiate most profusely at a wavelength of 10 μm, where an important atmospheric window exists.

This means that in practice the thermal-infrared observer wants everything in the FOV of the detector to either be very cold (so it radiates very little) or have extremely low emissivity (so it radiates very poorly), or both. Figure 4-9 shows the flux of thermal photons from a black body as a function of wavelength and the temperature of the black body. This should give some feel for the importance of thermal backgrounds for long wavelength work.

In practice we must live with the emission from the ambient temperature optics of the telescope. The primary, secondary and tertiary can have fresh reflectivities at 2 μm of about 99% (new silver is 99.4% reflective), so there will be a minimum of 3% net emissivity at ambient temperature. In addition, optical surfaces will become dusty and degrade with time. This might add another 3% total emission. Finally the primary mirror segment edges will add another 0.9% of almost diffuse background.

The blockage of the primary by the support structure is over 3%, but it is very discrete and a cold pupil mask in the scientific instrument should readily block its thermal radiation

We expect that with an excellent mirror coating and reasonable care in keeping the mirrors clean, the system emissivity (at ambient temperature) will be between 4 and 7%.

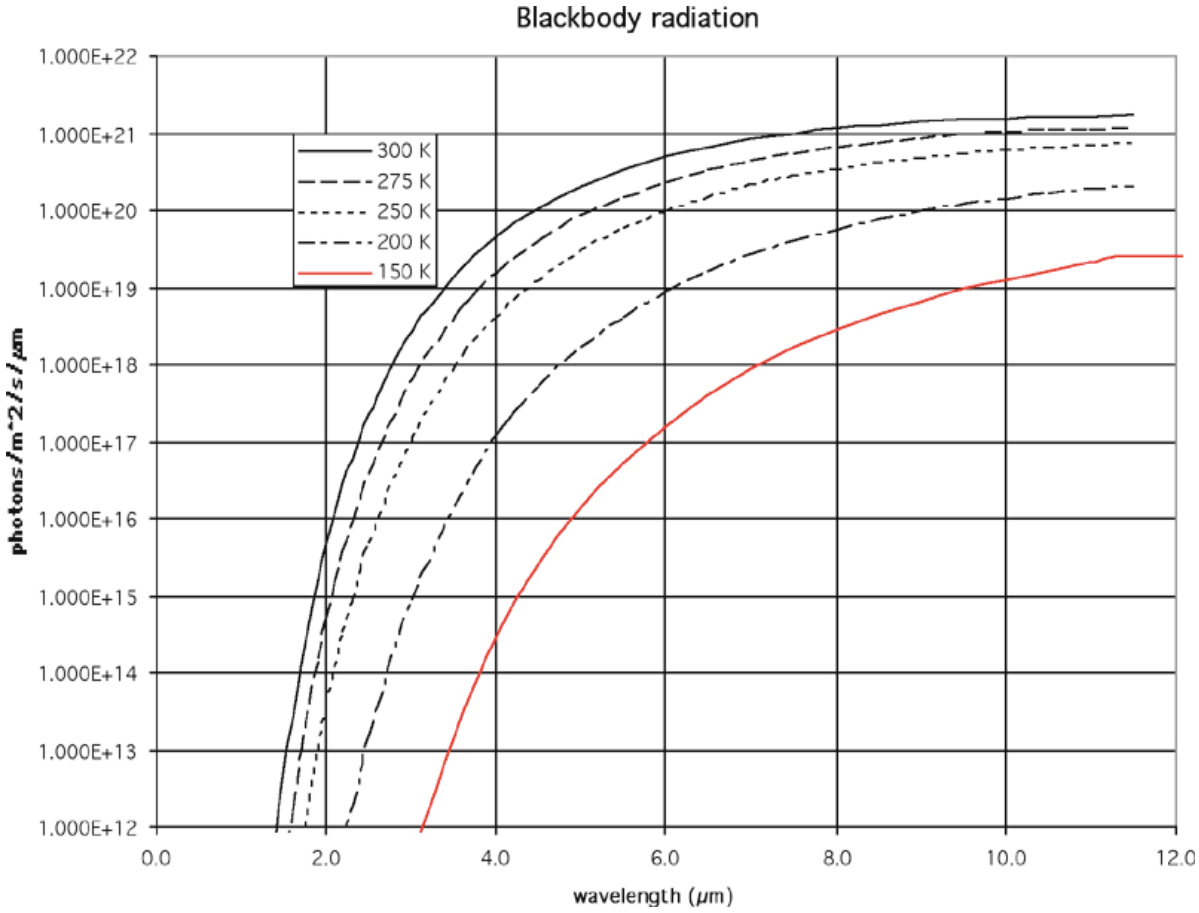


Figure 4-9. The flux from black bodies is shown as a function of wavelength for several different temperature black bodies. This highlights the importance of background temperature in producing background thermal emission.

REFERENCES

Faber, S.M. 1981. "Formulae Relevant to the Optical Performance of Telescopes." Keck Observatory Report No. 55.

Kuhlen, M. 2001. " The Tertiary Mirror Equations of Motion of an 'off-altitude-axis' Nasmyth Focus." CELT Report No. 22.

Schroeder, D. 2000. *Astronomical Optics*, Second Edition. Academic Press.

Thomas, N. and Jesse Wolfe. March 2000. "UV-Shifted Durable Silver Coating for Astronomical Mirrors." SPIE Proceedings **4003**, 312.

Chapter 5. Primary Mirror

5.1 Conceptual Design

The optical design of the primary mirror is described in Chapter 4. The primary will be a mosaic of individual hexagonal segments. An active control system uses sensors to periodically measure the relative piston/tip/tilt degrees of freedom of all segments, process the measurements, and send commands to three actuators per segment in order to stabilize these degrees of freedom in the face of gravitational and thermal perturbations. The active control system stabilizes the readings to a set of desired readings. The values of the desired readings are determined by a separate system called the Telescope Wavefront Sensor (Section 5.6). In this chapter we describe the segment design, fabrication, passive support, active control, assembly, alignment, and maintenance.

5.1.1 Segmentation Geometry

The segment size is defined by a segment radius (a) and segment thickness (h). The selection of a segment radius, and hence the number of segments, depends on a complex tradeoff of many costs. A larger segment radius increases the amount of asphericity required in the surface figure ($\sim a^2$), the gravity-induced deflections on a support ($\sim a^4/h^2$), the weight for handling ($\sim a^2 h$), and sensitivity to position errors in the array ($\sim a^2$). A smaller segment radius increases the number of active control actuators and sensors, the complexity of a control wavefront sensor, and the complexity of the alignment and control software.

The selection of the segment thickness is also a complex cost issue. A greater thickness will require larger required forces for intentional deformation during fabrication, greater cost of the blank material, greater thermal inertia in the telescope, and a greater mass for the support structure (the telescope). A smaller thickness will require more support points to maintain acceptably low deformations due to gravity. These in turn increase the reliability requirements.

At this early phase of the design we have not yet gathered accurate estimates of these costs and cost variations. This will be required to make final informed compromises. Based on our experience with the Keck telescopes and preliminary discussions with possible vendors, we have adopted for now a reference segment design. In the future this will select final values for the segment radius and thickness. In the meantime we have adopted a segment radius of a = 0.5 m and a segment thickness of h = 45 mm. The resulting array (Figure 5-1) contains 1080 segments. This is the most compact pattern (smallest circumscribing circle). A central subset of 19 segments is deleted from the array since the secondary mirror blocks light to them. The array has an area = 702 m^2**.**

We have the option of changing the positions of some peripheral segments to keep a closely circular periphery, and at the same time allow for convenient division of the array into subsets (full and partial clusters of 19 segments) for ease of handling. The cluster layout is shown in Figure 7-1.

Figure 5-1 shows the segment outlines as seen by a star, i. e., projected on the global X-Y plane. In this view the outlines are regular hexagons. Since the mirror is curved, the outline of a single segment viewed along its own normal is stretched in the global radial direction and is therefore a non-regular six-sided polygonal. For the outermost segments this elongation is 1.38%, and for a regular hexagon diameter of 1.0 m this is an increase of 13.8 mm. In the conceptual design phase we have used the approximation of a regular hexagon. In the next phase of design the elongation will be accurately included. It will affect many aspects to a small degree, including the required blank size, segment cutting, segment support, segment control, etc.

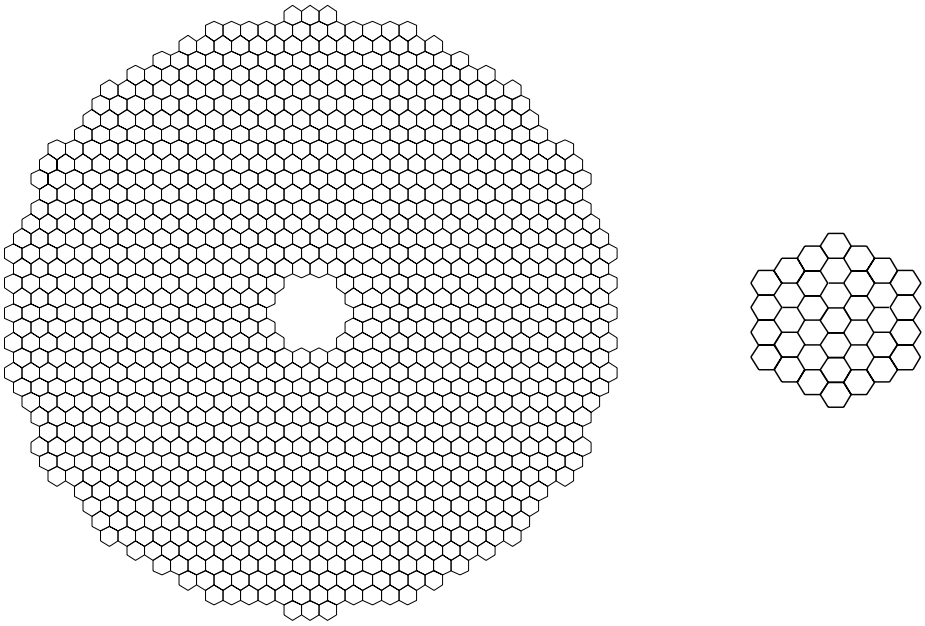


Figure 5-1. The primary mirror segmentation showing 1080 segments. The Keck primary mirror is shown to the same scale for comparison.

5.1.2 Rigid Body Control

A mirror or mirror segment is a flexible body, and its shape, as well as its position, are candidates for active control.

For Keck we chose to implement the segmented idea in its simplest form, in which the segment position is subject to active control, but its shape is set in fabrication and its passive support is designed to maintain this desired shape. Thus, only the six degrees of freedom that are needed to define a rigid body in space need to be controlled. With this decision the control system is greatly simplified. However, this does make more stringent demands on the fabrication and support of the segments.

The passive support of the Keck telescope segments was designed to maintain the surface figure of the segment, and also to control the three in-plane degrees of freedom. Only the piston/tip/tilt are actively controlled. Warping harness springs are used to manually change the segment figure; however, this only happens roughly once a year.

For CELT we will adopt a similar approach. However, the large number of segments for CELT will require remote control over the warping harness springs, if warping harnesses are needed. With remote control comes the potential for making real-time adjustments of low spatial order surface shape changes. The bandwidth for this control, however, can be much less than that of the active control system, since the surface deformations are small and induce slowly changing image effects. During the preliminary design phase the need for and efficacy of such a system will be studied.

For now we assume the same approach as used for the Keck telescope; the segments are treated as rigid elements with only six degrees of freedom to be controlled.

5.1.3 Active and Passive Degrees of Freedom

As a rigid body, each segment has six degrees of freedom. We characterize these as piston, tip and tilt (rotations about two axes in the segment plane), rotation about the normal to the segment, and radial and azimuthal motion of the segment center in the surface of the primary mirror. For the relatively slow focal ratio of CELT, only the first three motions (piston, tip, and tilt) require active control. The remaining three motions are assumed to be passively constrained by the rigidity of the support structure. Of the three "passive" motions, azimuthal motion has no optical effect.

The piston, tip, and tilt degrees of freedom are actively controlled with three actuators behind each segment. Each actuator is coupled to the mirror via a whiffletree (see Section 5.2.1). The three actuators are parallel to the segment surface normal at the center of the segment.

The tolerances on the other two degrees of freedom, rotation about the segment normal and radial translation of the segment center in the array, will require care to achieve. If installation errors or gravity- or temperature-induced changes in these two degrees of freedom are objectionable, we may require remote control of them (or of the surface errors, focus, and astigmatism induced by them). If this is the case, then the bandwidth for this control will be much smaller than that required by the active control system. The necessity and cost of remote control will be addressed in the preliminary design phase.

5.1.4 Active Control System

The active control system contains three major components: 1) a system of sensors to measure the relative positions of the segments, 2) a computer and global control algorithm that uses the sensor information to calculate the desired piston, tip, and tilt changes in the segments and to send motion commands, and 3) a system of actuators that re-position the segments in response to perturbations. We expect slower perturbations of the segments to occur from thermal changes and gravitational loading changes of the support structure. We expect faster perturbations to occur from wind forces on the segment. We emphasize again that the "active control system" is a stabilization system, stabilizing the sensor readings at externally determined values.

5.2 Segment Passive Support

Substantial progress was achieved on the segment passive support by consulting engineer Steve Gunnels (2001). Gunnels used finite element modeling to establish force and moment distributions that provide adequate segment support designs for both axial and lateral gravity loads. He also designed the hardware required to apply these loads. The Keck design required holes to be bored into the midplane of the glass segments, which was a complication. Both of Gunnels's current designs are accomplished without requiring holes in the segments. We list below the assumptions and requirements used for this design.

Table 5-1. Conceptual segment design and support requirements

Reference Segment Design

Regular hexagon side length = a = 0.5 m
Thickness = h = 45 mm
Density = 2.53 x 10^3 kg /m^3
Radius of curvature of front and back surfaces = 90 m
Material: Zerodur
Elastic modulus = E = 91 GPa
Poisson ratio = ν = 0.24
CTE = 1 x 10^{-8}/°C

Reference Support Requirements

Tolerance on surface deformations = 6 nm rms
Gravity: zenith angle range: -9.5 to 74.5 degrees
Range of control motion = 2 mm independently for each actuator
Temperature range: operations = 2 ± 8 °C
Temperature range: survival = – 20 to + 40 °C
Humidity Range = 0 to 100% condensing
Support envelope = hexagonal by < 30 cm
(i.e., entirely behind the segment, within 30 cm)
Weight < 20% of 74 kg = < 15 kg
Stiffness = > 60 Hz
Lifetime = infinite
Maintenance = zero
Attachment to actuators = simple and rapid
Cost < $2k
Vacuum compatible - yes
Allows simple and rapid segment removal

By decomposing the gravity load into components we divided the problem into separate axial support and lateral support problems. The Keck segments are supported by two separate subsystems that handle these separate force components. Gunnels similarly designed two subsystems for CELT. He also studied an integrated system, where a single support design addresses both components of gravity.

Several versions of the integrated design were studied, and all gave large (~ 1- 2 μm peak-to-valley) deflections. A separated-subsystem design yields peak-to-valley deformations of order 20 nm.

The goal of Gunnels's study was to develop a passive, robust, and economical support system that acts solely from the back surface using the axial support points. The goal was to limit the surface distortion of the baseline segment to about 5 nm root-mean-square (rms) due to rotating in the gravity field for zenith angles from 0 to 65 degrees.

In pursuit of this goal, about 50 finite-element models were created and used to calculate surface distortion under various support arrangements and load conditions. The models ranged in size up to 14,000 degrees of freedom with as many as 10 load cases each. In addition, numerous two-dimensional CAD layouts were made of potential mechanical designs. Gunnels wrote many brief intermediate reports describing the detailed status of these studies. A final report summarizes the studies and the final designs (Gunnels 2001).

5.2.1 Axial Support

Scaling from the Keck segment axial support, Gunnels assumed that an 18-point support would be adequate. A regular hexagonal segment has 12-fold symmetry (a right triangular region can be folded around the segment normal 12 times). Therefore, for an 18-point support there are1.5 supports for each of the 12 triangular regions. If we label the radial sides of the triangle A (long side) and B (short side), then we can define four topologies for the axial support: full support near the center of the triangle plus half support edge A, full support center plus half support edge B, two half supports edge A plus half support edge B, one half support edge A plus two half supports edge B. A half support is one shared by either adjacent triangular region. These topologies are shown in Figure 5-2.

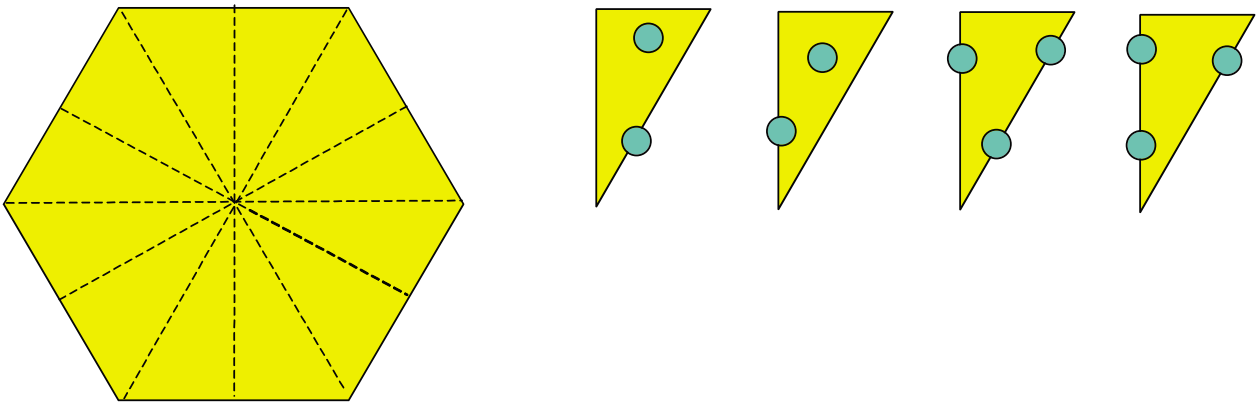


Figure 5-2 The four possible topologies for an 18 point axial support of a regular hexagon are shown. By symmetry, the hexagon can be divided into 12 equal right triangles, and each triangle must contain 1.5 supports.

All four topologies were analyzed, and the support locations optimized to minimize the peak-to-valley (p/v) surface distortion under 1 g zenith gravity. The rms for these surfaces was subsequently calculated. For the four topologies the p/v surface distortion ranged from 21.1 to 29.3 nm. The rms range was 5.06 to 7.11 nm. These rms values likely have a small error, since the rms calculation does not compensate for variations in mesh density in the finite element models. The two topologies with three half supports on the edges define mechanical support systems that are not symmetric with respect to the hexagonal shape of the segment. In addition, one of the symmetric topologies was revised allowing more droop of one corner (the point in the hexagon). This lowered its rms from 6.8 to 6.14 nm, while the p/v increased from 29.0 to 38.3 nm. The conclusion of this study was that lowest rms topology (5.06 nm, two half-points on A) requires an asymmetric whiffletree design, and the next lowest rms topology (6.14 nm, one half-point on A, one interior point) requires a symmetric whiffletree design, which Gunnels pursued in detail. In addition, if further reduction is desired, it may be practical to ion figure into the optical surface part of the expected zenith gravity deflections, thus significantly reducing the rms surface error averaged over zenith angle.

The conceptual design axial support hardware is shown in Figures 5-2 and 5-3. As shown, each of the three actuators supports a whiffletree that equally distributes the axial force via flexures to two triangular plates. The force applied to each triangular plate is then distributed to three points on the segment via the rod flexures. The 18 axial support forces are accurately defined, and the system is kinematic. The total weight of the system is 9.4 kg and a detailed cost analysis gives $695 per segment for hardware and assembly.

The true segment shapes will not be regular hexagons, and this needs to be accommodated in the final design of the axial segment support. For the Keck segments this was addressed by bonding weights to the periphery of each segment to approximate the load distribution of a regular hexagon (called the "superhexagon"). The axial support system was then designed for this larger regular hexagon, allowing the identical support to be used for all segments.

5.2.2 Lateral Support

If the segment is supported at one or more points on the back surface and lateral (horizon) gravity applied, a moment develops due to the axial eccentricity of the lateral force with respect to the center-of-gravity location in the central plane of the segment. Early analysis showed that, if the moment were only compensated by the whiffletrees, the p/v distortion of the segment would be hundreds of nm or more. In addition, through considerable optimization it was learned that the compensating lateral (shear) force should be applied over a relatively small area and that the moment compensation is best applied just outside, or beyond the footprint of the shear force application.

After numerous optimizing computer analyses (the results of which are described in the status reports), a lateral support geometry was arrived at which yielded very good results. The design uses a counterweight attached to the back of the segment with six silicon bonding pads (Figures 5-3 and 5-5). Bonding with six pads is locally over-constrained. However, the pad design is sufficiently compliant to keep surface distortion due to thermal and manufacturing effects to a negligible level. Finite element analysis of the design predicts a surface deformation from a 1 g lateral load that has a p/v of 10.5 nm and rms of 1.5 nm. The design hardware consists of a moment annular counter weight and extension tube, a moment force ring (that bonds to the segment), lateral force puck, disk flexure (that allows actuator-driven motion of the segment while carrying the lateral load), disk adapter fitting, and zenith counterweight. The total lateral support system weight is 12.2 kg and a detailed cost analysis results in $407 per segment for hardware and assembly. The distortion of the segment due to this support was calculated to be 10.5 nm p/v, with an rms of 1.8 nm.

The concept design of the axial and lateral supports meets our gravity distortion goals, is simple, and is cost effective. However, it does not quite meet our weight and envelope requirements. Further analysis is required of the disk flexure, axial rod flexures, and glued invar attachments to quantify other errors and determine whether the system is adequately robust.

In addition to creating the design, Gunnels studied the segment deformations due to wind loads. As expected, the deformations will be quite small compared to rigid body motions on the elastic passive support system. He completed a detailed cost estimate for designing, building, and mounting the passive support system on the segments. The total estimated cost per segment is $1100, substantially less than our goal of $2000/segment. Gunnels studied the fabrication and assembly error sensitivities, and these were used to establish the error budget values in Chapter 11. He calculated the surface deformation induced by moments around the pivot axes, and these were used in the warping analysis in Section 5.3.7. He also calculated the natural deformation frequencies of a segment on three actuators. The lowest frequency is ~ 350Hz.

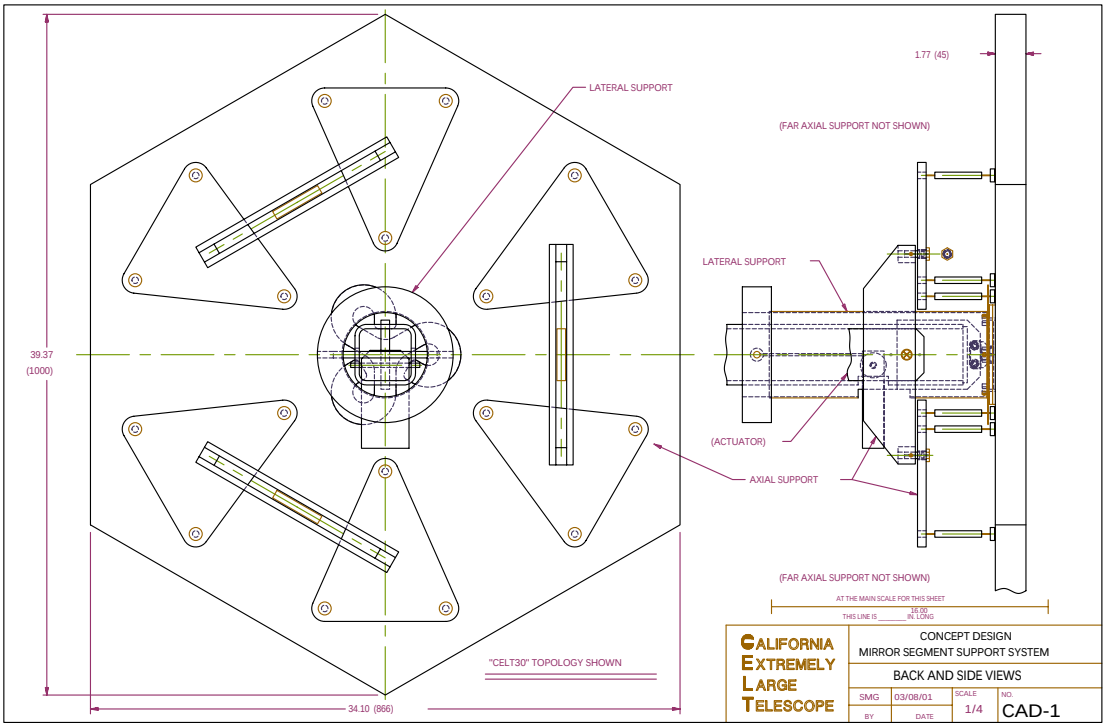


Figure 5-3. Design of axial and lateral segment support system.

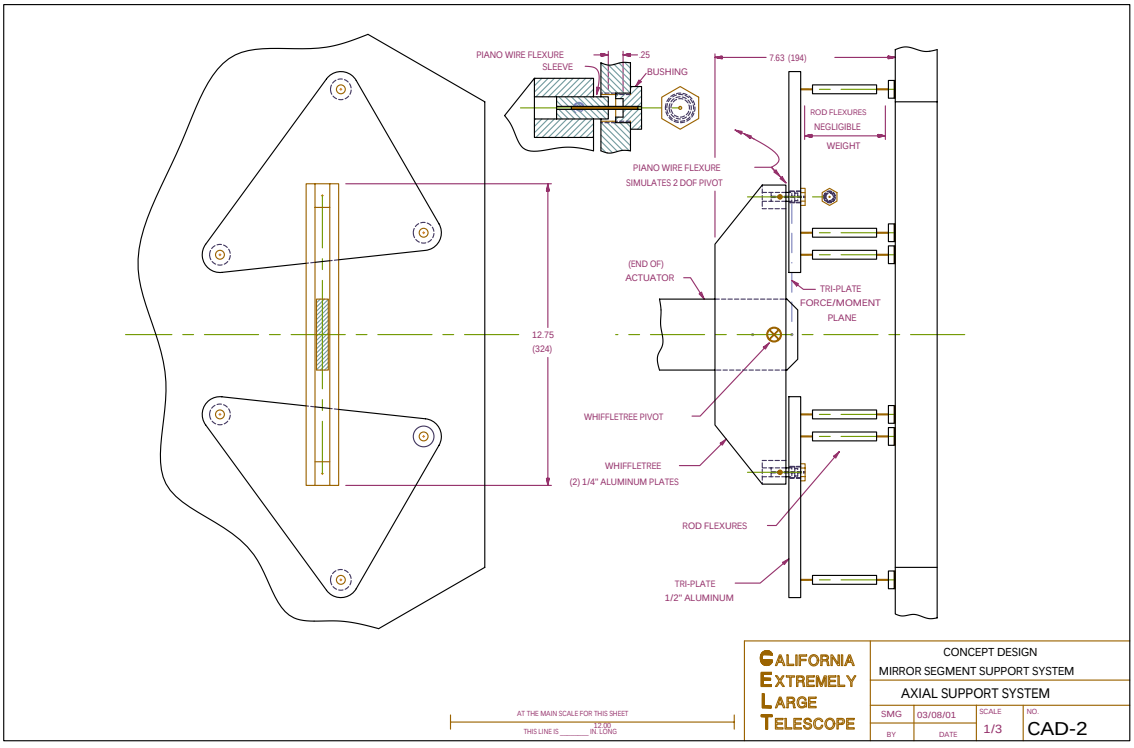


Figure 5-4. Design of axial segment support system.

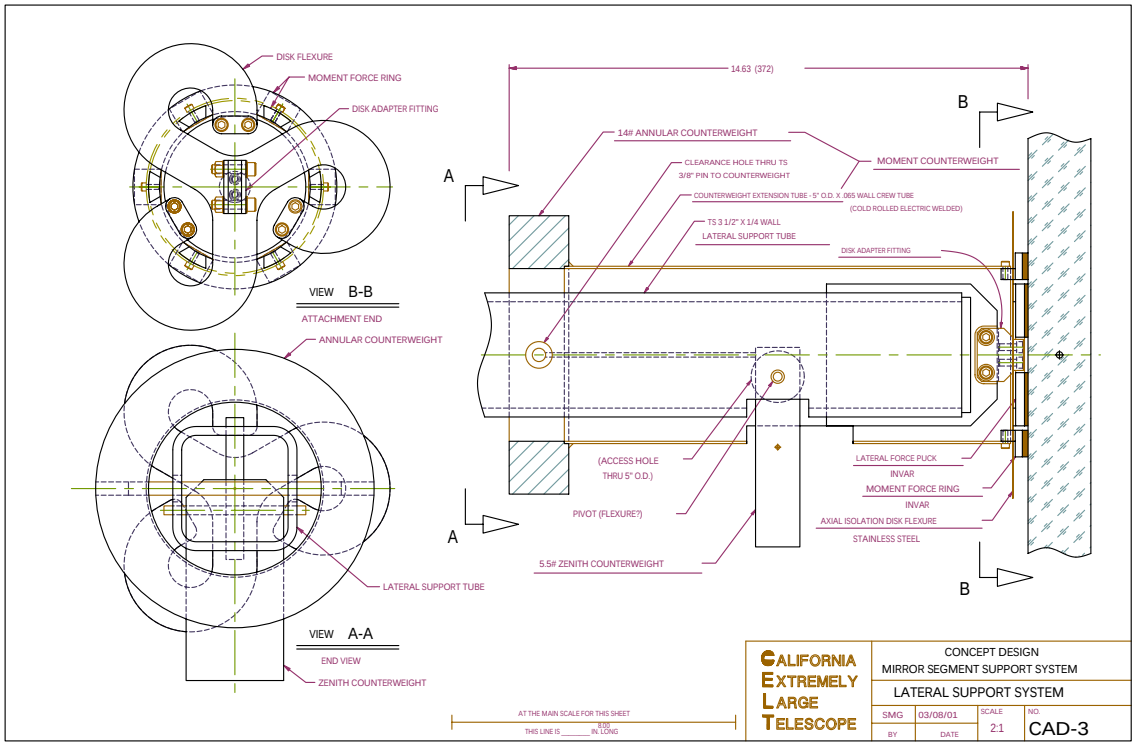


Figure 5-5. Design of lateral segment support system.

The performances of the axial and lateral support subsystems are now known and are a function of the telescope zenith angle. In principle, ion figuring can be used to remove the expected surface deformation at a particular zenith angle. We will get the best overall performance if the figuring is optimized for a typical zenith angle. During the preliminary design phase we will investigate this possibility. Input to this investigation will include the expected telescope operating zenith angle distribution, the error budget for AO-off and AO-on (including the AO deformable mirror response to the axial- and lateral-support-induced aberrations), surface testing errors, ion-figuring errors, and warping harness efficacy.

5.3 Segment Fabrication

Using our assumed reference design for the telescope optical design and primary mirror segmentation geometry, we can calculate the surface figure of each segment. We describe these in terms of an expansion in Zernike polynomials (Nelson, Gabor, Hunt, Lubliner, and Mast 1979). Each coefficient is a function of R = off-axis distance of the segment center, k = primary mirror radius of curvature, K = the conic constant, ρ = r/a (where r = segment radial coordinate and a = segment radius), and θ = the polar angle on a segment (= 0 along radial line). The segment surface in a coordinate system centered on the segment with the z-axis normal to the surface is given by

$$z(\rho,\theta) = \Sigma\, C_{ij}\, Z_{ij}\, (\rho,\theta). \qquad (5\text{-}1)$$

To lowest order in R/k, the dominant terms are

$$Z_{20} = 2\rho^2 - 1 \text{ (focus)} \qquad C_{20} = a^2/(4k) + Ka^2R^2/(4k^3) + \ldots \tag{5-2}$$

$$Z_{22} = \rho^2 \cos 2\theta \text{ (astigmatism)} \quad C_{22} = Ka^2 R^2/(4k^3) + \ldots \tag{5-3}$$

$$Z_{31} = (3\rho^3 - 2\rho)\cos\theta \text{ (coma)} \quad C_{31} = K a^3 R/(6k^3) + \ldots \tag{5-4}$$

For k = 90 m and a = 0.5 m these coefficients are plotted as a function of R in Figure 5-6.

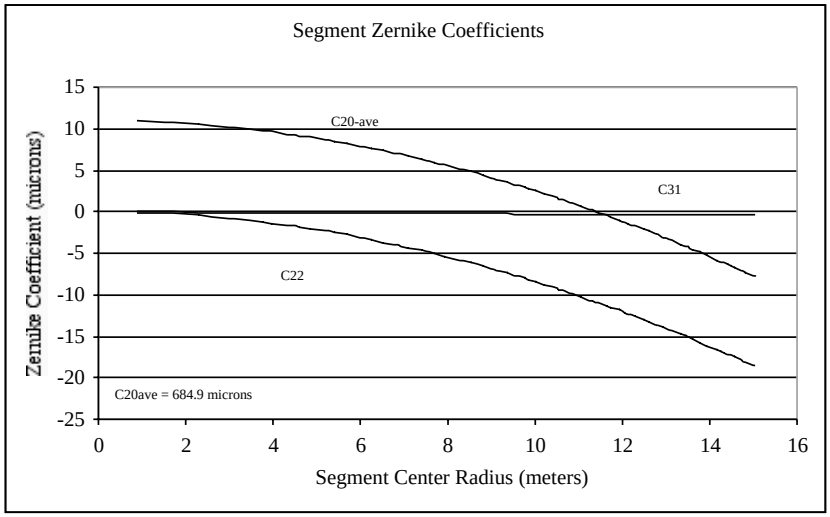


Figure 5-6. Segment surface Zernike Coefficients (μm) versus segment-center radius (m).

C_{20} varies over the range 694.4 to 675.7 μm. The outermost segment deviates most from a sphere; the coefficients for the outermost segment are given below.

Table 5-2. Zernike coefficients for the segment at the largest radius

	30-meter	**10-meter (Keck)**
a (meters)	0.5	0.9
R (meters)	15.0	4.68
C_{20} - C_{20} on axis (μm)	-18.8	-113
C_{22} (μm)	-18.6	-113
C_{31} (μm)	-0.4	-15

Our reference design for the mechanical shape of the segments assumes a 2.0 mm gap between segments. We assume the sides of the hexagonal prism are parallel to the local normal to the primary surface at the segment center. We also assume the back surface of the segment is a convex sphere with a radius = 91 m for all segments (the average front-side radius). With an assumed bevel of 1 mm, the non-reflecting gap is 4 mm. There are 3102 inter-segment gaps with a total fractional light-loss area of 0.0088. For comparison, the Keck segments are separated by 3 mm gaps and have 2 mm bevels; resulting in a fractional light-loss area = 0.0070.

5.3.1 Fabrication Process Overview

To date the only large segmented mirrors in use are those in the Keck telescopes and in the Hobby-Eberly Telescope (HET). The Gran Telescopio Canarias (GTC) is a segmented-mirror telescope under construction. The segment fabrication process will be different from that used for Keck; however, at this writing no segments have been produced. Because of this, and because the HET segments have spherical surfaces, we use the Keck segment fabrication as a basis for studying options for CELT. The Keck segment fabrication process is summarized in Table 5-3. Itek Optical Systems (Lexington, MA) and Tinsley Laboratories (Richmond, CA) polished the Keck segments. Eastman Kodak (Rochester, NY) did the ion figuring. The testing method used during or after each process step is indicated in Table 5-3.

Table 5-3. Segment fabrication process for Keck

Process Flow	Testing
Convex Side Polish	Spherometer
Stressed Mirror Grind and Polish	Profilometer
|	Autocollimation Test
Cut Hexagon and Drill Holes for Support	
|	Autocollimation Test
Ion Figure	
|	Autocollimation Test
Mount on Support and Install in Telescope	
|	Shack-Hartmann Test
Set Warping Harness (in telescope)	
|	Shack-Hartmann Test

Keck stressed mirror polishing included the following major steps:

1. Applying shear forces and moments at the edge of the circular blank via 24 levers to elastically deform the blank in order to induce a surface equal to (polished-sphere minus desired-hyperboloid). This required forces of 0 to 15 kg and moments of 26 to 46 kg-m per lever.
2. Polishing a sphere into the deformed blank.
3. Releasing the forces and moments; the blank elastically relaxed, and the sphere deformed into the desired hyperboloid.

A warping harness refers to a system of 30 leaf springs integrated with the axial support on the back of each segment. These are used to achieve the final segment figure based on in-telescope measurements using a Shack-Hartmann test and a bright star.

For CELT we propose the modification of this process outlined in Table 5-4. The details are described in the sections below.

Table 5-4. Proposed segment fabrication process for CELT

CELT proposed Process Flow	Testing
Convex Side Polish	Spherometer
Planetary Stressed-Mirror Grind and Polish	2-dimensional contact probe array
|	point-diffraction interferometry
Cut Hexagon	
|	point-diffraction interferometry
Ion Figure	
|	point-diffraction interferometry
Mount on Support and Install in Telescope	
|	Shack-Hartmann Test
Set Warping Harness (in telescope)	
|	Shack-Hartmann Test

5.3.2 Blank Material

Before a final selection of blank material is made we will study the costs and performance of various candidates. We are considering Zerodur (Schott Optical Technologies, Mainz, Germany), ULE fused silica (Corning Corporation, Canton, NY), aluminum, and composites.

The issues that will be studied in order to make a final decision include technology development required, internal stress that may cause unpredictable warping from cutting, polishability, support complexity, mass of support structure, stiffness-to-weight ratio, thermal stability, and material cost.

We have defined an initial set of specifications listed below, and include for comparison those used for the Keck telescopes.

Table 5-5. Blank specifications

	CELT	Keck
circular blank diameter (m)	1.05	1.9
meniscus uniform thickness (mm)	45	75
meniscus surface radius of curvature	91 m	35 m
stress birefringence mean less than	3 nm/cm	3 nm/cm
coefficient thermal expansion		
mean (rms) =	$< \pm 1.4 \times 10^{-7}$/°C	$< \pm 1 \times 10^{-7}$/°C
coefficient of thermal expansion axial gradient		
mean (rms) =	$< \pm 1.3 \times 10^{-8}$/°C/m	$< \pm 1.3 \times 10^{-8}$/°C/m
segment mass	74 kg	400 kg

For the conceptual design phase studies we have assumed the reference material to be Zerodur.

5.3.3 Generating and Grinding

As was done for the Keck segments, we will purchase the blanks with a uniform meniscus shape generated by the manufacturer. We will use loose abrasive grinding to remove the generating surface damage and to figure the surface in preparation for polishing. Whether this will be done one blank at a time or with multiple blanks simultaneously on a planetary grinder depends on a detailed cost comparison to be made in the next design phase.

5.3.4 Planetary Stressed-Mirror Polishing

The current reference design for CELT calls for a primary mirror with the following specifications:
radius of curvature (k) = 90 m
conic constant (K) = -1.00284
segment radius (a) = 0.5 m
off-axis distance (R) = 0 to 15 m

Nelson and Temple-Raston (1982) derived the polar monomial coefficients that describe the surface of an off-axis segment .

$$z(\rho,\theta) = \Sigma_{ij}\, \alpha_{ij}\, m_{ij}\, (\rho,\theta) \qquad (5\text{-}5)$$

$$m_{20} = \rho^2 \qquad \alpha_{20}^{\text{segment}} = a^2/(2k) + Ka^2R^2/(2k^3) + \ldots \qquad (5\text{-}6)$$

$$m_{22} = \rho^2 \cos 2\theta \qquad \alpha_{22}^{\text{segment}} = Ka^2 R^2/(4k^3) + \ldots \qquad (5\text{-}7)$$

$$m_{31} = \rho^3 \cos\theta \qquad \alpha_{31}^{\text{segment}} = K a^3 R/(2k^3) + \ldots \qquad (5\text{-}8)$$

$$m_{40} = \rho^4 \qquad \alpha_{40}^{\text{segment}} = (1 + K) a^4/(8k^3) + \ldots \qquad (5\text{-}9)$$

where the normalized radius $\rho = r/a$. The polar monomial coefficients are related to Zernike polynomial coefficients (Nelson, et al., 1979) as follows:

$$C_{20} = \alpha_{20} / 2 + \alpha_{40} / 2 + ... \qquad (5\text{-}10)$$

$$C_{22} = \alpha_{20} / 2 + ... \qquad (5\text{-}11)$$

$$C_{31} = \alpha_{31} / 2 + ... \qquad (5\text{-}12)$$

$$C_{40} = \alpha_{40} / 2 + ... \qquad (5\text{-}13)$$

Unlike the polar monomials, the Zernike polynomials have very useful orthogonal properties, and we use them in much of the remainder of this chapter.

In the stressed mirror polishing process (Lubliner and Nelson 1979; Nelson, et al., 1979) a sphere is polished into the elastically deformed blank (see Section 5.3.1). If the radius of the sphere is L, then

$$\alpha_{20}^{\text{sphere}} = a^2/(2L) \qquad (5\text{-}14)$$

$$\alpha_{40}^{\text{sphere}} = a^4/(8L^3) \tag{5-15}$$

The difference between the sphere and segment gives the desired stressing deformation

$$\alpha_{20}^{\text{str}} = a^2/(2L) - a^2/(2k) - Ka^2R^2/(2k^3) + \ldots \tag{5-16}$$

$$\alpha_{22}^{\text{str}} = - Ka^2 R^2/(4k^3) + \ldots \tag{5-17}$$

$$\alpha_{31}^{\text{str}} = - Ka^3 R /(2k^3) + \ldots \tag{5-18}$$

$$\alpha_{40}^{\text{str}} = a^4/(8L^3) - (1 + K) a^4/(8k^3) + \ldots \tag{5-19}$$

The average value of $\alpha_{20}^{\text{segment}}$ over the primary is

$$\alpha_{20}^{\text{segment}} \sim a^2/(2k) + Ka^2R_{max}^{\ 2}/(4k^3) \tag{5-20}$$

Setting L to match this average gives

$$L = [1/(k) + KR_{max}^{\ 2}/(2k^3)]^{-1} \tag{5-21}$$

$$\Rightarrow \qquad \alpha_{20}^{\text{str}} \sim - Ka^2 (R^2 - R_{max}^{\ 2}/2) / (2k^3) + \ldots \tag{5-22}$$

For the outermost segment of CELT (R = 15 m) the coefficients are

$$\begin{aligned} \alpha_{20}^{\text{str}} &= 18.8\ \mu m \\ \alpha_{22}^{\text{str}} &= 18.6\ \mu m \\ \alpha_{31}^{\text{str}} &= 1.2\ \mu m \\ \alpha_{40}^{\text{str}} &= 0.001\ \mu m \end{aligned} \tag{5-23}$$

An attractive method for introducing these terms is to use a modification of stressed mirror polishing that was used for Keck (the outermost Keck segment required $\alpha_{22}^{\text{segment}} = -113\ \mu m$, $\alpha_{31}^{\text{segment}} = 44.8\ \mu m$). The Keck segments were polished one at a time with a stressing fixture below the blank. The stressing loads were applied using gravity and lead weights. For CELT we plan to polish multiple segments simultaneously on a large planetary polisher with a stressing fixture above each blank.

The stressing fixture introduces the difference between the desired surface and the sphere being polished. To introduce the small C_{40} using the stressing fixture requires a uniform pressure on the blank. We limit the stressed mirror polishing to only forces and moments around the edge. We assume the small C_{40} will be created using ion figuring.

With a stressing fixture above and attached to the edges of each blank, the bottom surface of the blank is polished by the planetary polishing tool. When the stress forces and moments are released, the polished surface will be close to the asphere required.

For a reference design we assumed a 1.05 meter diameter circular blank and calculated the forces and moments required to deform it.

We first used analytic calculations to calculate the approximate sizes of forces and moments required to achieve the required aspheres. We then used a finite-element code to calculate more accurate values.

Our analytic estimate is based on Lubliner and Nelson (1979). The axial force density $V(\theta)$ and moment density $M(\theta)$ around the edge of a circular blank required to achieve given values of C_{20} and C_{22} are given below. For the calculation here we set C_{31} to zero. It is small and can be readily included in the stressed mirror polishing.

$$M(\theta) = M_0 + M_2 \cos(2\theta) \qquad V(\theta) = V_0 + V_2 \cos(2\theta) \qquad (5\text{-}24)$$

$$\text{and} \quad M_0 = (D/a^2)\, 2\,(1+\nu)\, 2\,\delta C_{20} \qquad V_0 = 0 \qquad (5\text{-}25)$$

$$M_2 = (D/a^2)\,(1-\nu)\, 2\,\delta C_{22} \qquad V_2 = M_2\,(2/a) \qquad (5\text{-}26)$$

$$\text{where} \quad D = Eh^3/[12(1-\nu^2)] \qquad (5\text{-}27)$$

E is the elastic modulus, ν is the Poisson ratio, h is the blank thickness.
The units are [M] = N, [V] = N/m, [E] = N/m^2, [h] = m.
For the proposed blanks: E = 9.06×10^{10} N/m^2, ν = 0.24, a = 0.525 m, h = 0.045 m.
The equations above yield D = 7.31×10^5 N-m, $M_0 = 1.32 \times 10^7\ \delta C_{20}$,
and $M_2 = 4.03 \times 10^6\ \delta C_{22}$.
For the outermost segment R = 15.0 m, $\delta C_{20} = -9.4$ μm
(assumes a planetary tool radius ~ C_{20} ave = 684.9 μm) $\delta C_{22} = -18.6$ μm.
These imply M_0 = 124 N M_2 = 75 N V_2 = 286 N/m.

This analytic estimate assumes the blank is a thin plate. For a more accurate model that includes the effects of the finite thickness of the blank and shear stresses, we have used a finite-element calculation using the program ANSYS. Since a finite number of levers are used in practice (24 for Keck), there are high spatial frequency local deformations at the edge of the blank where the forces and moments are applied. Most of these are removed when the circular blank is cut into a hexagon.

Cabak (2001) created detailed finite-element models to calculate from pure quadratic deformations these high spatial frequency deformations and the resulting deviations. These will be used to determine the optimal circular blank radius, the required density of levers, and the details of bonding the levers to the blank. His work to date models the stressing fixture to induce the dominant aberrations, C_{20} and C_{22}. He modeled 25% of the blank and included 12 levers to apply the desired moments and forces to the edge of the blank (corresponding to 48 levers for the full blank). Each lever is bonded to the back of the segment and has a radial width of 25 mm and azimuthally covers half the circumference. With the loads required for the outermost CELT segment this results in surface stresses of about 2.8×10^6 Pa (400 psi), a safe loading for the Zerodur blanks. Cabak's Model 31 includes 114,700 nodes; 99,972 elements; 341,979 degrees of freedom; and runs in about 1900 seconds. The global surface deformation closely agrees with the analytic prediction. A best-fitting quadratic surface (C_{20} and C_{22}) was then subtracted in order to see the very small deviations. These are shown in the contour map of Figure 5-7. The range over the full circular blank radius is 123 nm. The range over a circle of radius 0.5 meters (the circle enclosing a hexagon with side length = 0.5 meters) is 74 nm. The rms average over a regular hexagonal area is 7.0 nm.

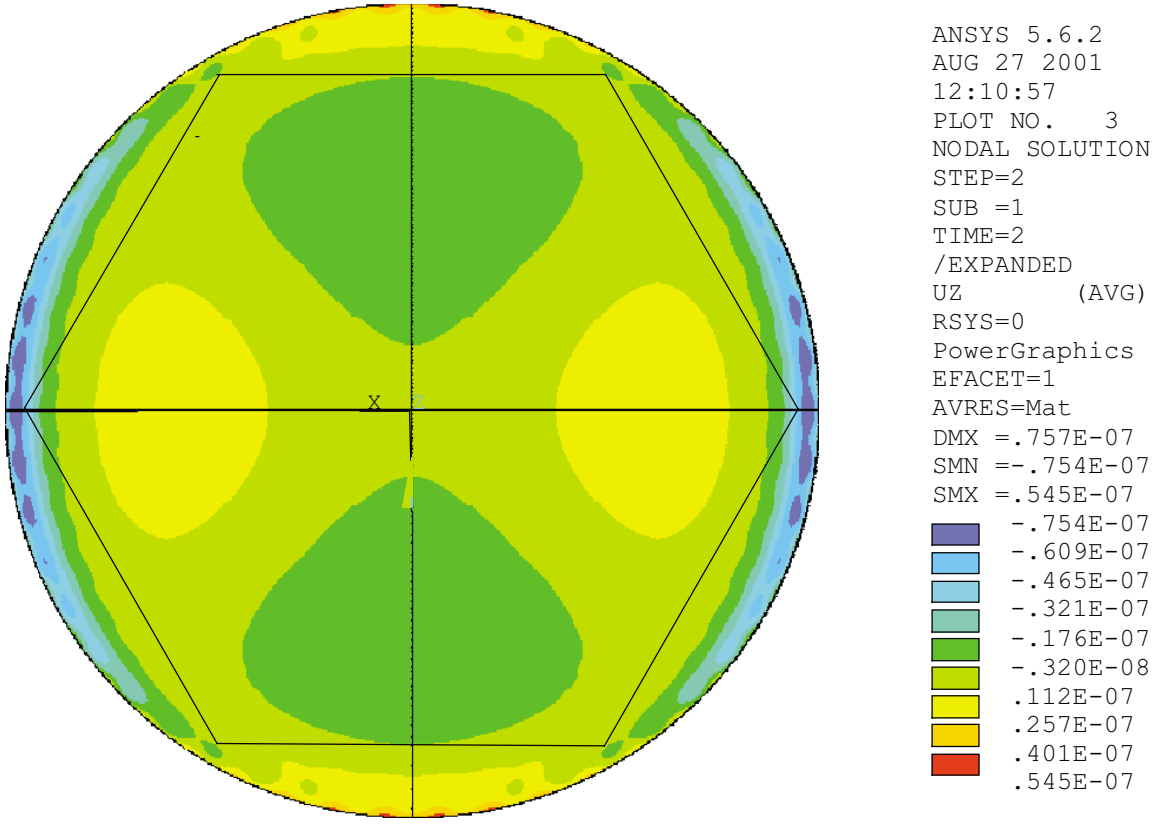


Figure 5-7. Finite-element predicted deformation non-quadratic residual surface (units are meters).

Figure 5-8 shows the axial stress component in the blank. The maximum stress is about 400 psi (2.8 x 10^6 N/m^2), which agrees well with an analytic estimate and is several times smaller than the breaking tensile strength of Zerodur.

We have also begun to design an engineering-model stressing fixture. Laiterman (2001) describes a 48-lever design where springs are used to apply two forces between the lever end and a fixture ring to induce moments and forces to the blank (Figures 5-9 and 5-10). The spring lengths are adjusted using a "robot" that is placed and rides on the fixture ring. It is motor-driven from lever to lever and uses two load cells to adjust the spring forces. Since the system is self-equilibrating, multiple iterations around the levers will be required to achieve a specific force and moment distribution. Figure 5-10 shows a design for the capturing, unlocking, adjusting, and locking of each spring length. The robot is removed for polishing and testing.

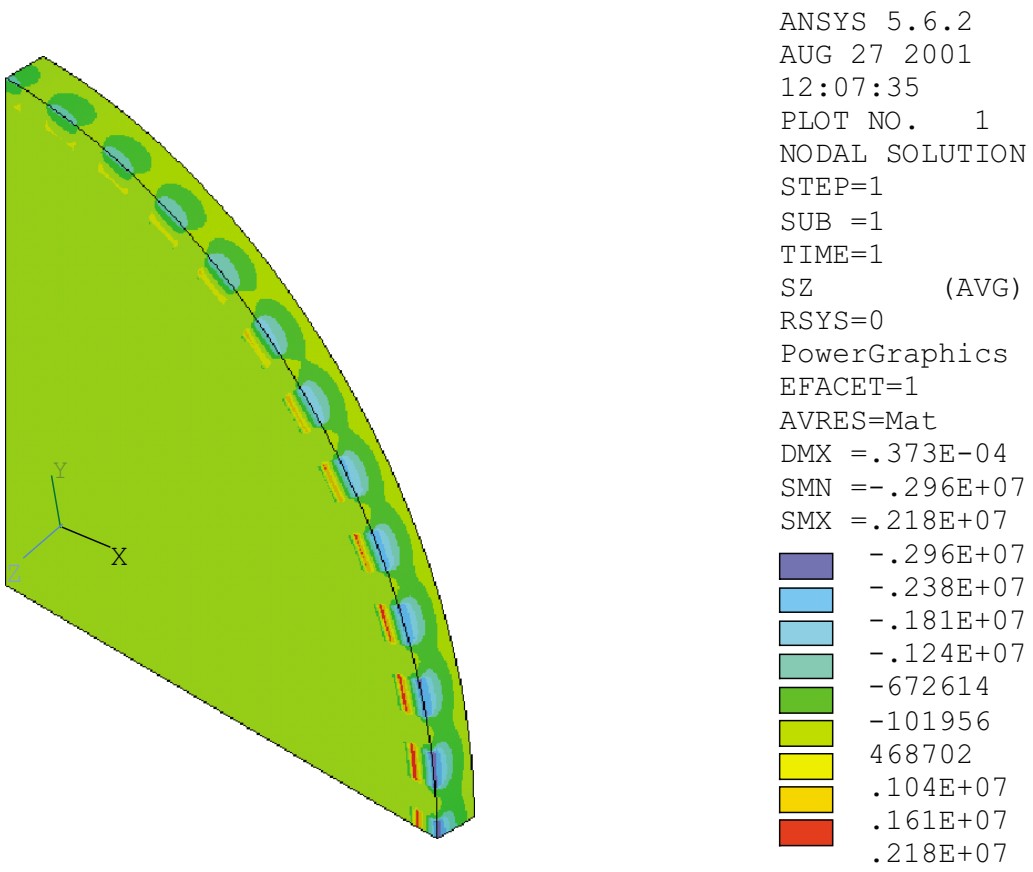


Figure 5-8. Finite-element predicted axial stress distribution (units are Pa).

We have begun to address the practical issues of planetary polishing. Today there exist three large planetary polishing machines in the United States with diameters of 3.7, 4.1, and 4.6 m. In the preliminary design phase we will study the applicability, availability, and costs of these polishers for the CELT segments. The very large number of segments may warrant the fabrication of a new polisher specifically designed for this fabrication program.

5.3.5 Cutting

The circular mirrors for the Keck segments were cut into hexagons using a diamond saw. The option of using a water knife (a high velocity water stream loaded with abrasive) was tested and found to give a cut surface that was too rough. Most of the roughness occurred at the outgoing edge of the cut where the stream wandered more. Since the CELT segments are thinner (45 mm) than those for Keck (75 mm) and since water knife technology has improved, we will test this option again for both performance and cost.

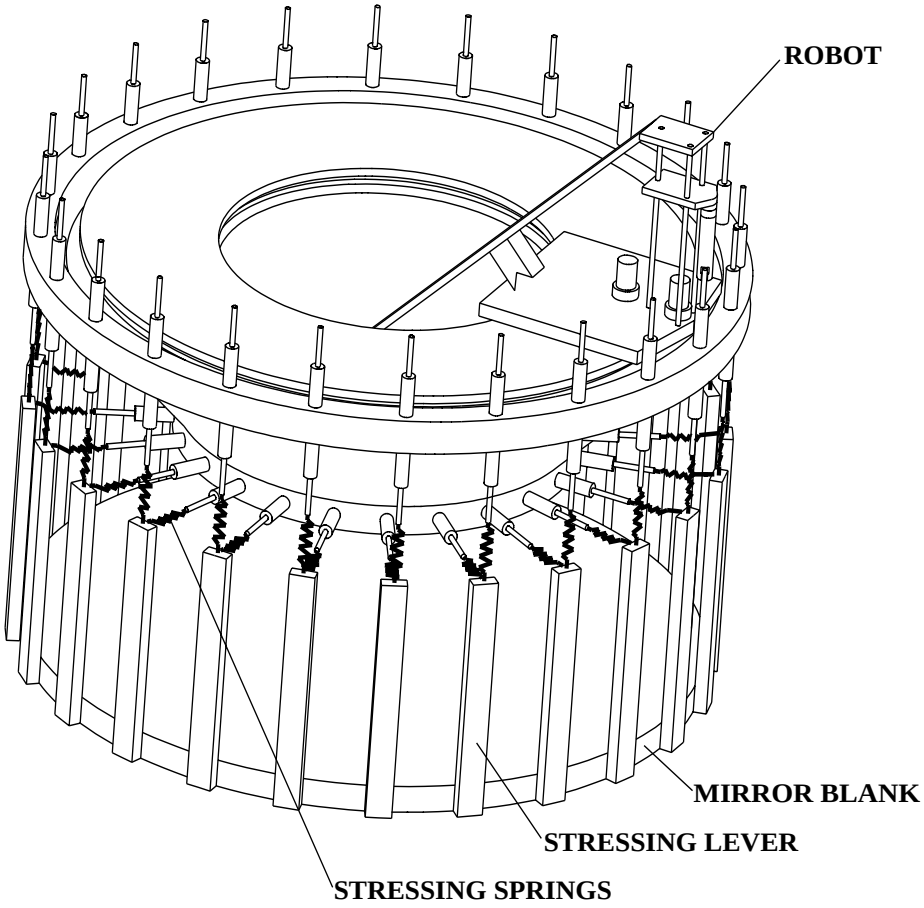


Figure 5-9. Stressing fixture showing the blank and levers bonded to the back at the edge.

5.3.6 Ion Figuring

Ion figuring played a critical role in refining the figuring of the segments for the Keck telescopes. Assuming the same areal density of residual material to be removed for the CELT segments implies removing $30^2/(2 \times 10^2) = 4.5$ times the material removed from both Keck primary mirrors. If the same ion figuring facility were used, then two years of continuous operation would be required. For an estimated segment fabrication time of five years this is not a problem. Ion figuring technology and experience have advanced since the Keck segments were figured, and we will pursue using these advances to make the ion-figuring process even more feasible and cost effective. Potential avenues include improved ion guns, multiple ion guns, oblique (rather than normal) incident beams, more energetic beams, different element beams, and multiple stations operating in parallel.

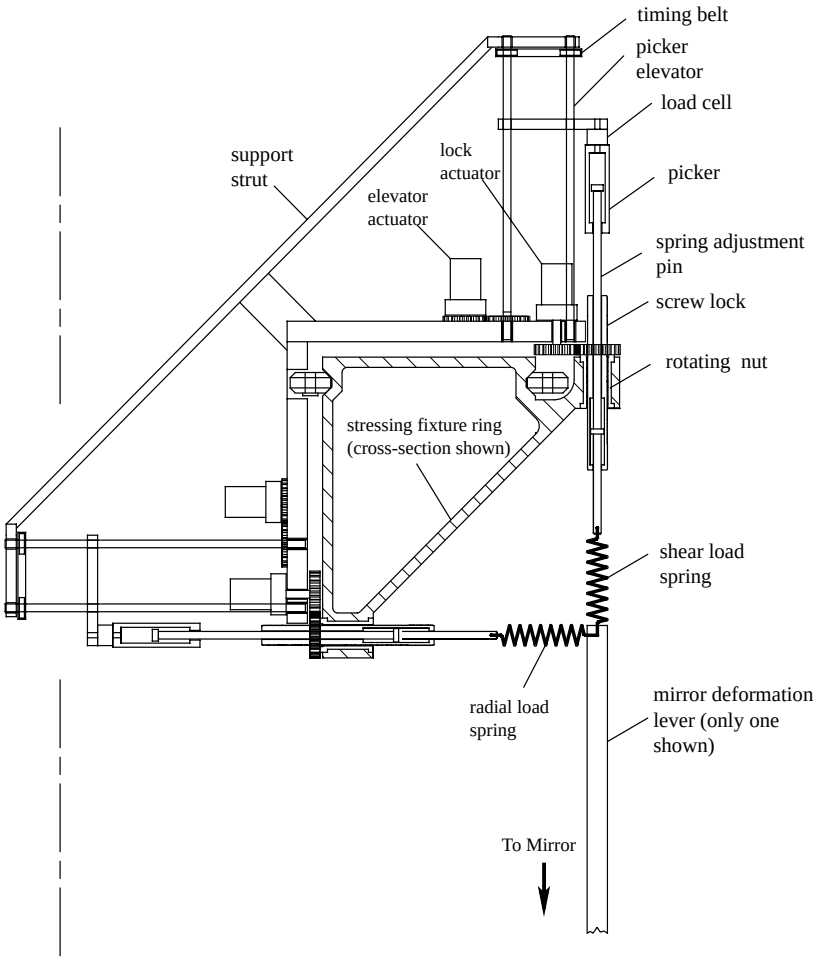


Figure 5-10. The mechanism on the robot used to adjust the spring loading of the stressing fixture.

5.3.7 Warping Harnesses

There are multiple sources of errors that might be introduced between the end of the fabrication production processing and the use of the segments in the telescope. Warping harnessses are used to correct these errors. The final figure of the Keck primary mirror segments is achieved using a set of thirty leaf springs on the back of each segment to apply forces through the segment support to deform the surface small amounts. Each segment surface is measured in the telescope using a bright star and a Shack-Hartmann camera. These measurements are used to determine the desired reading on a strain gauge bonded to each leaf spring. The design of the system and its use are described by Panoskaltsis, et al., (1987); Zaslawsky, et al., (1989); Mast (1989); and in many Keck Observatory Technical Notes.

Each leaf spring applies a moment about one of the flexural degrees of freedom of the segment axial support structure. The moments applied to the whiffletrees provide a significant improvement of the low-order Zernikes. Reduction of the lower order aberrations necessarily introduces some higher spatial frequency aberrations. A measure of the power of the technique is the ratio of residual rms surface error divided by the initial rms surface error.

Adjusting the Keck manual system is awkward and time-consuming. If warping harnesses are necessary, a remote-controlled system will be required for CELT. The exact number and range of forces will be determined by the errors expected from the fabrication, support, and installation of the segments.

We studied the efficacy of an initial design of a warping harness system for the CELT segments (Mast and Nelson 2001). We assumed a regular hexagonal segment with the reference segment size, side length = a = 0.5 meters. We assume the segment axial support is a set of three whiffletrees, each supporting the segment at six points, with a total of 18 support points.

We assumed the axial support system designed and analyzed by Gunnels (2001). Figure 5-11 shows schematically the layout and labeling of the support points and pivots.

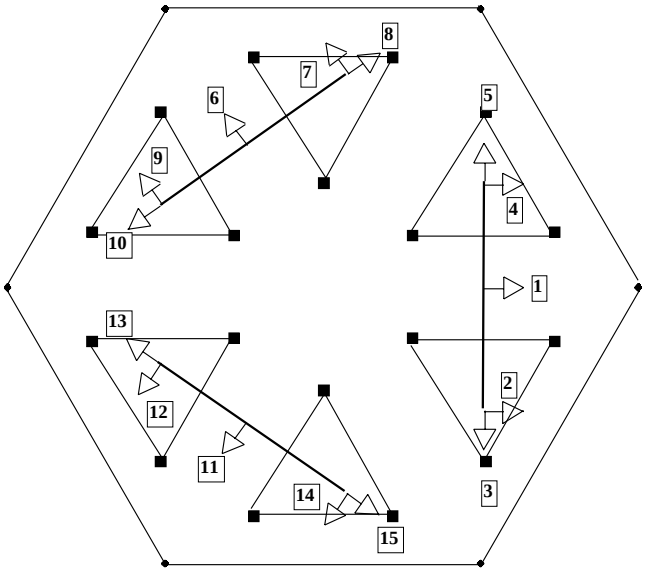


Figure 5-11. Layout of the support points and pivots used for warping harness analysis.

Included in Gunnels's studies were three runs to calculate the surface deformations induced by a 1.0 N-m moment about a single pivot for pivot numbers 1, 2, and 3. Each yields a surface deformation of δz heights at 2299 x, y points; upper surface node z-translations (nm) in the model. These nodes approximately, but not exactly, sample the surface in equal areas. The mean, range, and rms about the mean of each surface are reflected below in Table 5-6.

Table 5-6. Surface deformations induced by 1 nm moment about the three basic pivots

Pivot	mean (nm)	range (nm)	rms about mean (nm)
1	0.0	157.6	28.5
2	-0.2	99.6	20.0
3	-0.6	114.7	24.0

We reflected about the x-axis the surfaces from moments 2 and 3 to create deformation surfaces for moments 4 and 5, and then used rotations by 120 and 240 degrees to create surfaces for moments 6 through 15. Thus a set of 15 basis surfaces, one for each pivot moment and each with 2299 x, y, δz, coordinates, was created (Mast and Nelson, 2000). Each surface is induced by 1.0 N-m about a single pivot.

We then wrote a warping harness least-squares fitting program that fits any surface to a linear sum of the 15 basis surfaces, and used it to fit input surfaces that were single low-order Zernike polynomials of amplitude = 1.0 nm.

Table 5-7 below shows the input surface rms, the residual to the fit, and the "reduction factor" (final/ input rms surface height) for each low-order Zernike,

Table 5-7. Fitting a single Zernike deformation to the 15 basis surfaces

			CELT			**Keck (KOR 163)**		
polar		**cartesian**	**input rms**	**rms res**	**reduct. factor**	**input rms**	**rms res**	**reduct. factor**
C_{2-2}	$\rho^2 \sin 2\theta$	$2xy$	0.352	0.107	**0.304**	0.342	0.003	**0.009**
C_{20}	$2\rho^2 - 1$	$2x^2 + 2y^2 - 1$	0.526	0.176	**0.334**	0.516	0.024	**0.047**
C_{22}	$\rho^2 \cos 2\theta$	$x^2 - y^2$	0.352	0.107	**0.304**	0.342	0.003	**0.009**
C_{3-3}	$\rho^3 \sin 3\theta$	$3x^2 y - y^3$	0.249	0.021	**0.083**	0.303	0.012	**0.040**
C_{3-1}	$(3\rho^3 - 2\rho)\sin\theta$	$3x^2 y + 3y^3 - 2y$	0.319	0.212	**0.664**	0.317	0.034	**0.107**
C_{31}	$(3\rho^3 - 2\rho)\cos\theta$	$3x^3 + 3xy^2 - 2x$	0.319	0.212	**0.664**	0.317	0.034	**0.107**
C_{33}	$\rho^3 \cos 3\theta$	$x^3 - 3xy^2$	0.317	0.041	**0.131**	0.237	0.000	**0.000**

The "input rms" values for CELT and Keck difference is due to both the orientation for the Keck calculations being rotated 30 degrees from CELT's, and the node distribution for the CELT input not being uniformly distributed. The "reduction factors" for CELT are much closer to 1.0. Thus, the CELT design is much less effective than that for Keck. Table 5-8 shows the comparable quantities for the Keck warping harnesses from Panoskaltsis, et al., (1987).

Table 5-8 shows comparison of the CELT and Keck segment and support parameters.

Table 5-8. Support parameters: CELT and Keck

	CELT	**Keck**
segment radius = a (m)	0.50	0.90
segment thickness = h (m)	0.045	0.075
a^4 / h^2	31	117
support points	18	36
pivot moments	15	30

We conclude from Table 5-8 that both the smaller number of attachment points and the smaller number of moments of the CELT segments greatly reduces the efficacy of this warping harness design. In the preliminary design phase we will study the efficacy of increasing the supports to 36 per segment, and alternative designs for warping the segment.

Resources during the preliminary design phase will also be directed to understanding a long-standing problem with the Keck warping harnesses. On average the Keck warping harness systems improve the rms surface error of the segments by a factor of two. This is substantially less than the predicted factor of order ten. Whether this discrepancy arises from software or hardware limitations is not known. Before an effective warping harness design for CELT can be credible, the Keck warping harness problem needs to be resolved.

5.3.8 Segment Figure Testing

The critical importance of measuring the surface during fabrication motivates the creation of a detailed plan and a technology development program for segment testing. We want to measure segments in both circular and hexagonal shapes. We will need to design supports for both shapes that can be used during the fabrication process. The deformations induced by these supports will be minimized in the design, measured to high precision, and used in the test data analysis. Ideally, for the final measurements the support will be identical to that used in the telescope.

For the fabrication of the Keck segments we used a profilometer during the polish-test cycles. The profilometer was a linear array of LVDTs on a graphite-epoxy beam. It was calibrated on a reference surface, then set on the segment surface and manually or electronically clocked in azimuth. For in-process testing of the CELT segments we envision using a two-dimensional array of LVDTs, eliminating the need for clocking. This array will receive from above either a reference surface or the segment (since in PSMP the optical surface is face down). Readout, correction for support-induced deformations, reference surface calibrations, and a fit to polynomials will be scripted to provide rapid feedback to the polishing process.

For surface measurements at the conclusion of polishing and after ion figuring we will use an interferometric test. This test must accommodate the large aspheric departure while maintaining an accuracy of 15 nm rms surface error or better. For segmented optics, the relative segment curvatures must be controlled accurately (~ 2 parts in 10^5).

There are several possible interferometric tests that could be used to make these measurements. To achieve the highest measurement accuracy and most efficient processing, we assume phase-shifting interferometry is used for each option.

Major sources of error in interferometric tests include:
1. metrology errors in the test configuration
2. turbulence in the light path
3. vibration-induced position changes between optical elements
4. differential optical path errors due to the non-common path of the measurement and reference wavefronts
5. distortion of the aspheric mirror coordinate system in the detector plane
6. magnification calibration errors

We have made an initial study of the advantages and disadvantages of three options for interferometric testing. To significantly reduce the sensitivity to sources 1 and 2, all options use converging lenses to reduce beam distances to less than about 6 meters.

Mast, Nelson, and Sommargren (2000) describe an initial study of interferometric testing options. Since the outermost segment surface will contain about 20 μm of astigmatism, the fringe density is expected to be high. Three options for reducing the fringe density were considered in the above study: computer generated holograms, test configuration geometry of the final segment use (null test), and a tilted lens to create a canceling astigmatic wavefront.

Under the direction of Gary Sommargren, a group at Livermore has pursued the design of this test. They have considerable experience using a point diffraction test to measure X-ray optics surfaces to a precision of about 1 nanometer. Based on their experience at LLNL, they have concluded that, although the fringe density directly from the segments is high, it is still low enough to be measured without reduction. They have developed and analyzed a test that uses a phase shifting diffraction interferometry (PSDI) test to measure the segment (Sommargren, Phillion, Seppala, and Lerner 2001). The proposed test is conceptually simple and can be used to measure all segments in the CELT primary.

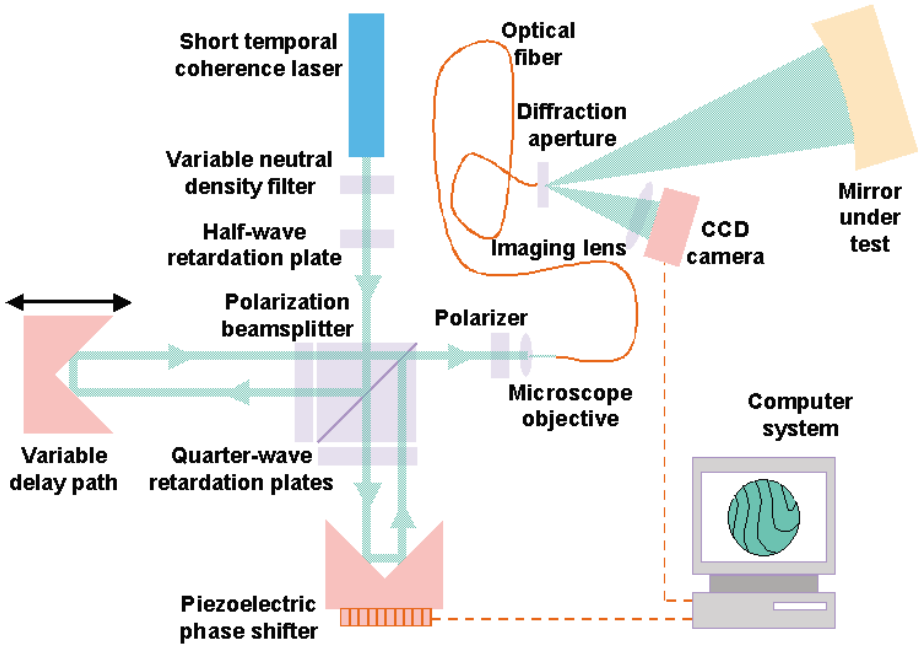


Figure 5-12a. Layout of point diffraction test.

The PSDI test configuration is illustrated in Figure 5-12a and has the following elements. Not shown is the condensing lens used to shorten the physical length of the test. This is illustrated in Figure 5-12b.

- Laser (λ = 532 nm).
- Fiber to provide the diffracting aperture.

- Polarization beam splitter: sends beams to interferometer reference and to converging lens.

- Converging lens: This lens is used to reduce the size of the test from about 90 m to about 7 m so the test can fit on a single optical table. The lens is placed about 5 m from the end of the fiber. Depending on the segment under test, the distance from the lens to the segment is varied from 75 to 2515 mm. Its position is accurately controlled using a distance-measuring interferometer. There are options for the design of this lens: singlet with asphere surface, triplet with spherical surfaces, condense/compensator lens combination, etc. A study of cost, performance, and tolerances needs to be made before selecting a design.

- Camera/Imaging Lens: A four-element camera was designed to image the surface on the detector (magnification ~ 1/74) and to minimize shearing effects and distortion.

- CCD detector: A camera used at LLNL for this type of test uses a 1024 x 1024 CCD. With that CCD, the fringe density in testing the CELT segments directly would be close to, but not above, the Nyquist frequency. If necessary, it is certainly feasible to acquire or build a camera with more pixels.

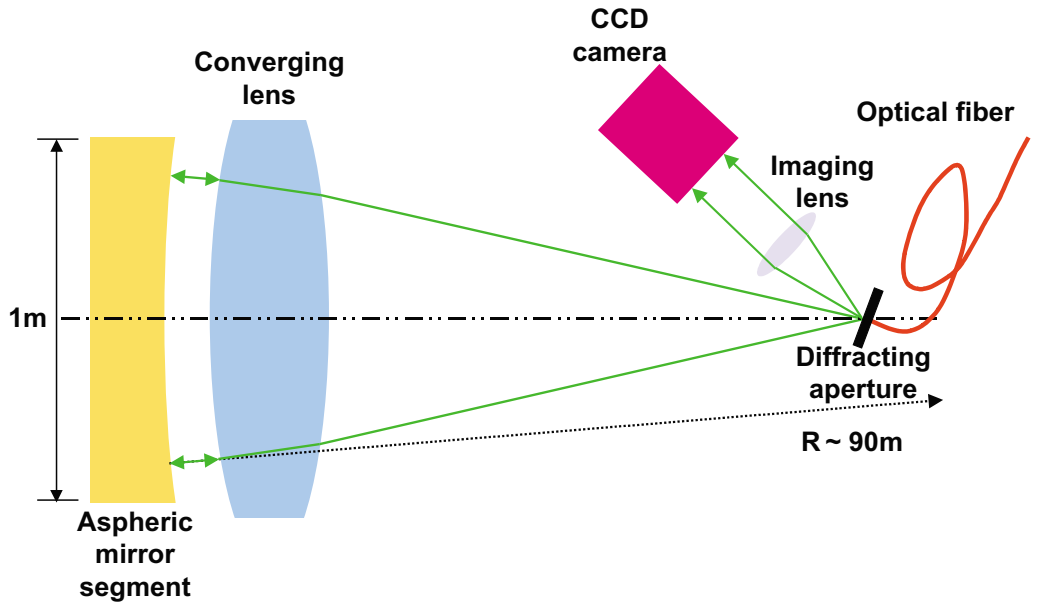


Figure 5-12b. Converging lens used to shorten the length of the interferometer.

There are three critical issues in the test:
1. differential optical path errors due to non-common path of the measurement and reference wavefronts for an aspheric segment
2. distortion of the local mirror coordinate system in the detector plane
3. magnification calibration

The Livermore group has addressed these issues both in the design and in proposed calibration procedures that use additional fixtures:
1. spherical test mirror (radius nominally = 91 m)

2. distortion calibration grid
3. magnification calibration mask

They have written software to simulate the test, including segment surfaces, coordinate systems, measurement configuration, etc. Holographic testing (Burge 2001) is also a promising approach that will be considered more carefully in the next phase.

5.4 Segment Active Control

Properly supported, CELT segments can be treated as rigid objects. Their position can be described by six parameters. Three of these (in-surface motions) can be adequately controlled passively. Three actuators attached to the cluster support structure will actively control the other three (piston, tip, tilt).

Perturbations Requiring Correction

There are both temperature- and gravity-induced changes in the axial positions of the segment supports. The control system moves the actuators to correct for these support changes. These changes are large enough to clearly require active control, yet they change sufficiently slowly that a relatively low system bandwidth is adequate.

As the elevation angle of the telescope changes, the changing direction of gravity deforms the mirror cell (see Chapter 7). In moving from horizon to zenith the deformation is dominated by a parabolic sagging of the cell. We expect the outer segment supports will move in piston about 2 mm relative to those of the inner segments. Thus the rms surface error will change by about 1 mm in going from zenith to horizon. The maximum rate of piston change is about 140 nm/second. The segments will tilt, and without active control the rms image radius would grow to about 100 arcsec in diameter.

As the temperature of the steel support structure changes, the mirror support positions also change. Over an expected operating range of about 10°C and without active control, the relative segment heights change by about 150 μm. In addition, temperature variations within the structure of 1°C will produce comparable distortions.

Control Hardware

Four hardware systems are required for the active control.

1. Sensors: We will use both displacement sensors on the edges of the segments and a Shack-Hartmann wavefront sensor observing a star. The information from both subsystems will be combined and used simultaneously in the control algorithm.

2. Actuators: three per segment.

3. Processor: The processor calculates the control matrix from the system geometry and then for each control cycle performs the matrix multiplication required to calculate the desired actuator motions.

4. Alignment Camera: Using starlight this instrument establishes the desired sensor readings. These desired sensor readings result in segment piston/tip/tilt degrees of freedom that geometrically stack the images from the segments and accurately phase the segment surfaces. The lookup table of the desired sensor readings may be a mild function of zenith angle.

All four systems must be designed in concert since the requirements for each depend strongly on the others. Our work to date has addressed primarily the displacement sensors (Section 5.4.2) and actuators (Section 5.4.3). The alignment camera is critical to the success of the system, and since it is a complex optical instrument we will address its design early in the next design phase. The wavefront sensor subsystem is also potentially complex; however, it is less specific to CELT and its detailed design can be delayed. Although the control matrices are large (potentially ~3300 by ~8500), the processor performance requirements will not require CELT-specific technology development. We will delay the design and specification of the processor hardware to a later phase of the design program.

The number of sensors and actuators for Keck and CELT are as follows.

Table 5-9. Sensors and actuators for Keck and CELT

	Keck	**CELT**
Number of inter-segment edges	84	3102
Number of displacement sensors	168	6204
Number of actuators	108	3240

The number of actuators and displacement sensors for CELT are about thirty times those required for Keck. Each displacement sensor for Keck cost roughly $5000 and each actuator roughly $7000. Given the large quantities, it is imperative to dramatically reduce the costs of these elements.

5.4.1 Control Algorithm

Three actuators located at the vertices of an equilateral triangle will adjust the piston, tip, and tilt of each segment. The control algorithm calculates a set of actuator moves in response to a set of sensor readings.

Two sensor subsystems, displacement sensors, and a wavefront sensor will measure the relative orientations of the segments. Each displacement sensor measures locally the relative height of two adjacent segments. There will be two sensors along each inter-segment edge for a total 6204 displacement sensors. The wavefront sensor subsystem uses a star to generate x, y centroid coordinates from each subaperture. We have not yet defined the number of subapertures, but we need enough to measure at least focus, tip, and tilt of the array; and the subapertures should be large enough for accurate centroids on very faint stars. The full vector of sensor readings contains two parts, displacement-sensor displacements and wavefront-sensor centroids.

There are many more sensor readings than actuators, and thus the system is highly over-determined. The full sensor vector will be used in a chi-square fit to determine the best estimate of the piston, tip, and tilt of each segment. Based on these, commands will be sent to the actuators to drive the sensor readings to externally determined *desired sensor readings.* This read-and-correct cycle will be repeated approximately ten times per second. (Keck actuators are updated two times per second [Jared, et al.,1990; Cohen, et al.,1990].)

The vector of sensor reading changes, $\boldsymbol{\delta s}$, that results from a vector of actuator motions, $\boldsymbol{\delta p}$, is defined by a matrix **A**.

$$\boldsymbol{\delta s = A\, \delta p} \qquad (5\text{-}28)$$

The elements of **A** will be calculated from the geometry of the segments, the locations of the actuators and displacement sensors, and the mapping of the wavefront-sensor apertures to the primary mirror. The control requires solving an over-determined set of equations. The result of a linear chi-square fit is a pseudo-inverse matrix B. The matrix B depends only on the matrix A, and given the system geometry, it will be calculated and stored. For each control cycle, the optimum actuator motions, $\boldsymbol{\delta p^*}$, are calculated using B

$$\boldsymbol{\delta p^* = B\,(s - s_d)} \qquad (5\text{-}29)$$

where s_d is the vector of desired sensor readings

and

$$\boldsymbol{\delta p}_{applied} = g\, \boldsymbol{\delta p^*} \qquad (5\text{-}30)$$

where g is the control system gain ($g \leq 1$) and $\boldsymbol{\delta p}_{applied}$ is the command vector sent to the actuators.

The desired sensor readings (s_d) are determined by measuring optically the positions of the individual segment images in the focal plane and the piston errors in the surface of the array. These measurements will be made using the alignment camera. This instrument for the CELT active control system has not been designed; it will be similar to that used for Keck (Chanan 1994a, 1994b, 1998, 1999, 2000). The vector s_d is the vector of sensor readings that results, as closely as possible, in all segment images being coalesced and all relative piston errors being zeroed.

Focus Degree of Freedom

The displacement edge sensors described below are sensitive to both relative vertical motion of adjacent segments edges, and also to relative tilt or a dihedral angle between adjacent segments. In this sense they are the same as the Keck sensors. The system of sensors measures all $3N_{segments} - 3$ degrees of freedom (all relative piston, tip, and tilt of all segments) of the array. One of these degrees of freedom is special ("focus mode"). It is the lowest spatial frequency mode, consists of the same dihedral angle between all adjacent segment edges, induces the same reading in all sensors, and is effectively a global curvature change in the array.

For a dihedral angle, ω, the maximum surface at the outer edge of the array is

$$S_{max} = (\sqrt{3}a/2)\omega(n/2)(n+1) \qquad (5\text{-}31)$$

where a = 0.5 m and n = 18 for the reference design. The resulting 100% image diameter, $\theta(100) = 26.2\,\omega$. If this global focus is removed by axial motion of the secondary then the radius of curvature of the array will not match that of the segments, and a segment-to-segment scallop results. The 100% image diameter induced by this scalloping is

$$\theta(100) = 4\sqrt{3}\,(2a/D)^2\,\omega\,(n/2)(n+1). \qquad (5\text{-}32)$$

We can define a characteristic length, L_{eff} , that relates the sensor reading, R, response to pure piston, δz, and pure dihedral angle ω; $L_{eff} = (\delta z/\omega)\,(R_\omega / R_{\delta z})$. The value of L_{eff} depends on the details of the sensor design. For Keck L_{eff} = 55 mm, and for the reference CELT sensor described below L_{eff} = 57

mm. The CELT sensor reading response to a pure dihedral angle is ωL_{eff}. Using this value the 100% image diameter due to focus mode is 0.095 arcsec/nm of sensor reading, and the 100% image diameter due to scalloping is 0.0024 arcsec/nm of sensor reading.

5.4.2 Sensors

Mechanical

For CELT we propose a mechanical sensor design that will be substantially less expensive than that used for Keck. The sensor capacitors will be bonded to the surfaces of the segments inside the inter-segment gaps (Figure 5-12). For a reference design we assume the mechanical sensor is a pattern of capacitor plates, leads, and ground planes. The deposited leads wrap below and onto the back of the segment where a bonding wire connects it to a small card with a sense pre-amplifier and cable connectors for the drive and sense electronics.

Although the sensor plates may be installed/applied with rather loose positional tolerances (~ ± 150 μm), once installed/applied we will measure the offset against a reference sensor to much higher accuracy (> 1μm). This will allow segment installation to provide initial segment alignment to ~ 1 μm once the control loop is closed using the offset values.

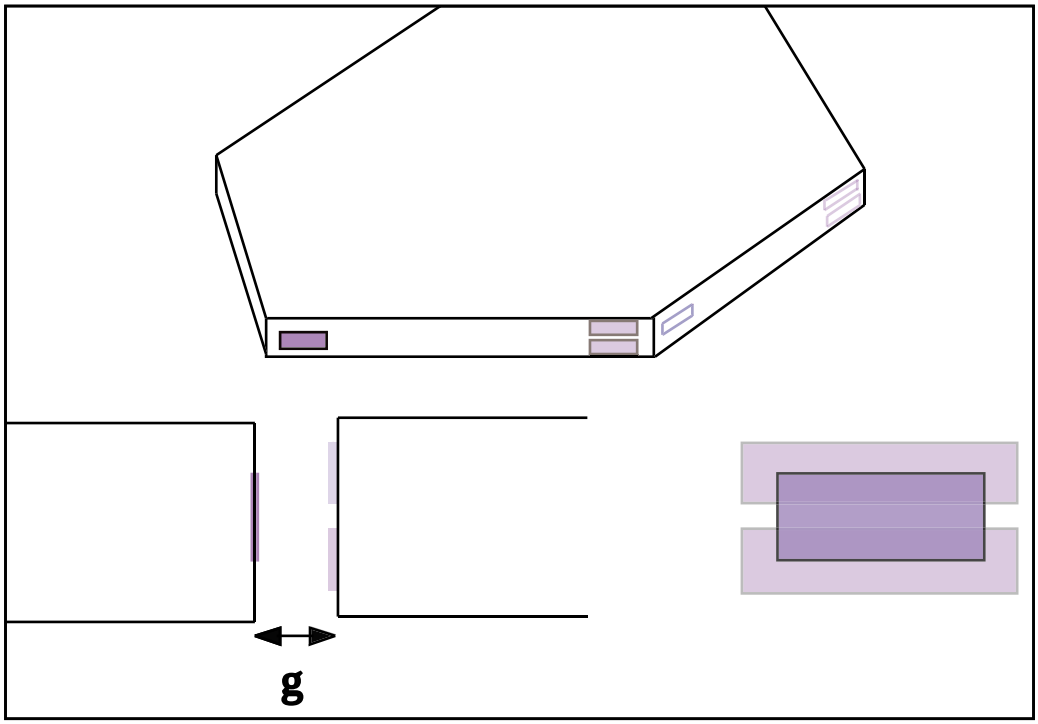


Figure 5-13. Illustration showing overall sensor positions.

Figure 5-13 shows a preliminary layout of the sensing and drive capacitors. We have called out some of the dimensions to be determined. The drive capacitor plates (c, d) are on one segment, and the sense plate (b, k) is on the segment across the gap (g). The basic signal that measures the vertical (z) displacement (δz) of one segment with respect to its neighbor is generated by the difference in capacitance of the sense capacitor with respect to the drive_1 and drive_2 capacitors. A displacement δz induces an output

$$\Delta Q = V_{drive}\ \varepsilon\ 2\ \delta z\ k\ /\ g \qquad (5\text{-}33)$$

where ε is the dielectric constant of air.

For a dihedral angle, ω, between segments (and no relative displacement δz) the sensor output charge will be [using $\Delta C = 2\ \varepsilon\ A\delta g/g^2$, $A = k(b-e)/2$, $\delta g = k\ \omega\ (b+e)/2$ and neglecting terms of $(e/b)^2$]

$$\Delta Q = V_{drive}\ \varepsilon\ b^2\ k\ \omega\ /\ 4g^2 \qquad (5\text{-}34)$$

The sensor output is proportional to the total charge, and thus the sensor is sensitive to a particular combination of δz and ω. This combination is used in defining the values of the control **A** matrix.

In addition to being sensitive to δz and ω, the drive/sense capacitances are also sensitive to a change in the gap size (as well as the dielectric constant of the air). Changes in the gap size will be induced by deformation of the telescope mirror cell as the telescope changes elevation angle and by changes in the temperature of the steel cell.

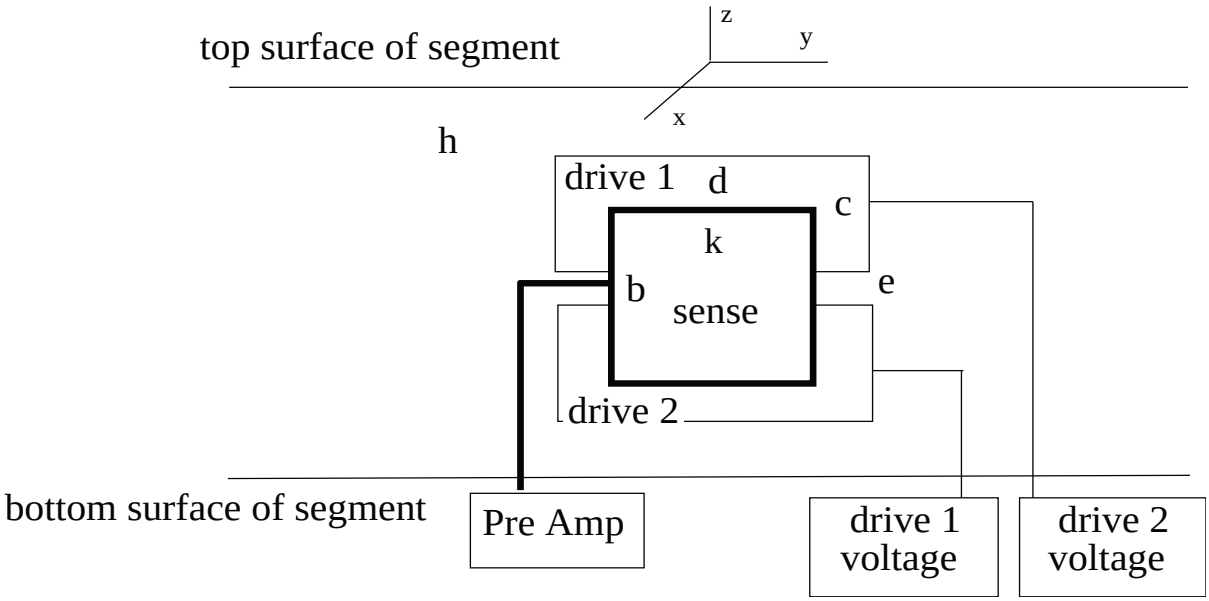


Figure 5-14. Defining sensor capacitor plate dimensions: b, k (sense); d, c (drive 1 and drive 2); e (gap between drives 1 and 2); h (blank thickness).

If the steel cell changes temperature by 10°C, then the gaps between segments will change by about 100 μm, or 5% change in the total gap. If the sensor is operating around the physical null, then this creates a small fractional error in the sensor reading and only slightly increases the control cycles required to iterate to the desired null. However, in practice large offsets may be required, of order 200 μm; and a 5% change in this large offset greatly exceeds the allowed errors.

The gravity-induced deformation of the cell also causes differential motions between segments and changes the gaps. In the Keck telescopes this causes the center positions of the segments to move radially up to 0.37 mm and rotate about the segment center up to 0.00025 radians (Mast 1987). This

would induce changes that are a large fraction of the 2 mm gap. An important goal of the CELT mirror cell design and cluster support structure will be to keep the gravity-induced gap changes to less than 100 μm.

To reduce the sensitivity of the sensor to gap changes we will (as with Keck) add a small difference between the two drive voltage plates until the electrical output is near null, when the segment height differences is zero. As a result, gap changes will still effect the sensor sensitivity, but will not shift the location of the null.

To reduce the adverse consequences of these two effects we will also measure the gap by shifting the drive voltage #1 with respect to #2 by 90 degrees instead of the 180 degrees used for Keck. Two synchronous detection circuits will then measure both the sum and difference in the capacitances. The sum provides the measure of the gap width needed to correct the difference measurement.

With Keck the sensor gains are known to 0.1%. This is useful for various calibration tasks, and this is a goal for CELT.

Cross-gap coupling

In principle, the drive voltages will originate from one segment, and the sense signal can be read on the other segment. However, this will result in a large ground loop that may be sensitive to noise pickup. To minimize this sensitivity we prefer to have both the drive and sense electronics on the same side of the gap. This could be achieved by coupling a small cable across the gap to carry the drive voltages or sense signal as is done for Keck. However, this would require connecting and disconnecting the cable for segment removal and installation. An attractive alternative is to have both the sense and drive electronics on the same side of the gap (same segment) and capacitively couple the signals across the gap. This will require additional capacitors. A detailed design, an engineering model, and a prototype of the complete capacitor system will be made during the preliminary design phase.

Electronics

The CELT edge sensor electronics will be similar to those used for Keck. The cost of 16-bit ADCs are now less than they were when Keck was designed, and will be used in place of the 12-bit ADCs used for Keck. This will provide greater offset and operating ranges. The major difference will be a relative shift of the phase of the drive and detection signals. With a second synchronous detection circuit this will allow detection of both the difference and sum of the capacitance changes.

Engineering-model sensors

We are building and testing engineering-model sensors of the design described above. We have begun by using the unmodified Keck electronics. In the preliminary design phase we will modify the electronics to read both the sum and difference in the capacitances, and develop the cross-gap coupling capacitors.

Sensor noise

Gary Chanan has coded the control matrix and calculated the response to sensor noise. Then using a singular value decomposition (SVD), Chanan calculated the singular values and orthogonal control modes of the control matrix. Each mode is a vector of actuator positions. The lower order vectors correspond approximately to the shapes of low-order Zernike polynomials. The table below gives the error "multiplier," M, for the first 30 modes where the modes have been ordered by decreasing singular value size. Each multiplier is the rms actuator value for that mode (in μm) divided by the rms sensor noise (μm). The rms actuator value is nearly identical to the rms surface error, S_{rms} and we can write

$$M = S_{rms} / \sigma_s \quad (5\text{-}35)$$

Chanan defines a dimensionless parameter, α_s, to describe the ratio of the focus mode sensitivity to normal mode sensitivity that depends on the details of the sensor design.

$$\alpha_s = b^2/(12\ t\ g) \quad (5\text{-}36)$$

where b is defined above in Figure 5-13, t = the radius of the actuator triangle, g = the sensor gap. For a baseline segment / capacitor design (t = 255 mm, g = 2 mm) Chanan considered three examples for α_s and calculated the following multipliers, M (Table 5-10 and Figure 5-15a).

The lowest singular value decomposition (SVD) mode corresponds to focus mode, a shape that has the same change in dihedral angle for all edge pairs. This mode is only sensed by the dihedral angle change-response of the edge sensors. All other modes are also directly sensed by segment height changes at the sensors. Tilt sensitivity is directly proportional to α_s.

Table 5-10. Edge sensor noise multipliers

	α_s =	**0.05**	**0.1**	**0.15**
	b =	**17.5 mm**	**24.7 mm**	**30.3 mm**
	SVD mode	**M**	**M**	**M**
	1	24.656	12.694	8.764
	2	8.774	8.198	7.541
	3	8.774	8.198	7.541
	4	3.908	3.656	3.360
	5	3.761	3.531	3.252
	6	3.287	3.014	2.685
	7	3.287	3.014	2.685
	8	2.296	2.145	1.964
	9	2.296	2.145	1.964
	10	1.780	1.657	1.500
	11	1.780	1.657	1.500
	12	1.551	1.444	1.317
	13	1.550	1.444	1.317
	14	1.550	1.442	1.305
	15	1.149	1.067	0.972
	16	1.117	1.051	0.963
	17	1.094	1.020	0.930
	18	1.045	0.986	0.906
	19	1.032	0.964	0.876
	20	1.032	0.964	0.876
	21	0.851	0.792	0.722
	22	0.851	0.792	0.722
	23	0.780	0.732	0.670
	24	0.780	0.732	0.670
	25	0.744	0.697	0.634
	26	0.744	0.697	0.634
	27	0.665	0.618	0.563
	28	0.665	0.618	0.563
	29	0.623	0.585	0.535
	30	0.572	0.541	0.499
rms of all 3240 modes		29.3	19.5	16.1
rms without mode #1		15.8	14.8	13.5

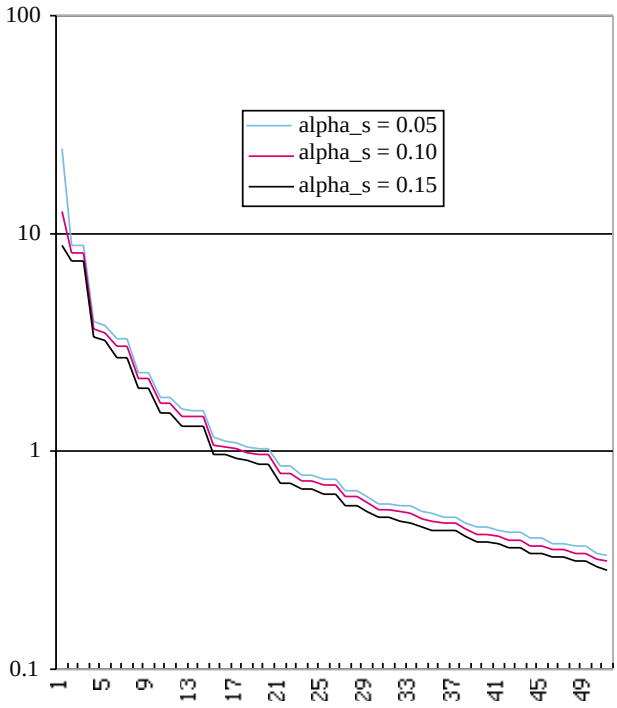


Figure 5-15a. Shows the multipliers for the first 50 modes. The first mode is focus mode. Only for this mode is the dependence on α_s strong. The multiplier values, M, for mode #1 are proportional to $1/\alpha_s$.

Figure 5-15b shows the residual error multiplier after completely correcting the first modes. This multiplier times the rms sensor noise gives the rms actuator value. The multipliers are again shown for the values of α given in Figure 5-15a, and the multipliers for Keck are shown in comparison.

The geometric rms radius due to focus mode is given by

$$k \sigma / (a \sqrt{N_s}) \qquad (5\text{-}37)$$

where k = 8.764, σ = the rms sensor noise, a = segment radius = 0.5, and N_s = number of sensors. If the global focus is approximately removed by axial motion of the secondary, the residual rms image radius is 0.0028 arsec/nm due to the scalloping of the wavefront.

Although all modes are measured by the edge capacitive sensor system, the lowest order modes have larger noise multipliers. To improve the precision of our knowledge of these modes or to add redundancy to the measurements, we may add information from the telescope wavefront sensor (WFS) to the control. The low-order modes change slowly, and thus the bandwidth required for a WFS could be lower than that for the edge sensors. The telescope WFS will use a guide star to measure tip and tilt (for telescope guiding), wavefront focus (for telescope focus), coma (for secondary mirror tilt and decenter), and additional lower order modes (for segment control). During the preliminary design phase we will study the efficacy of using this additional information for the segment control.

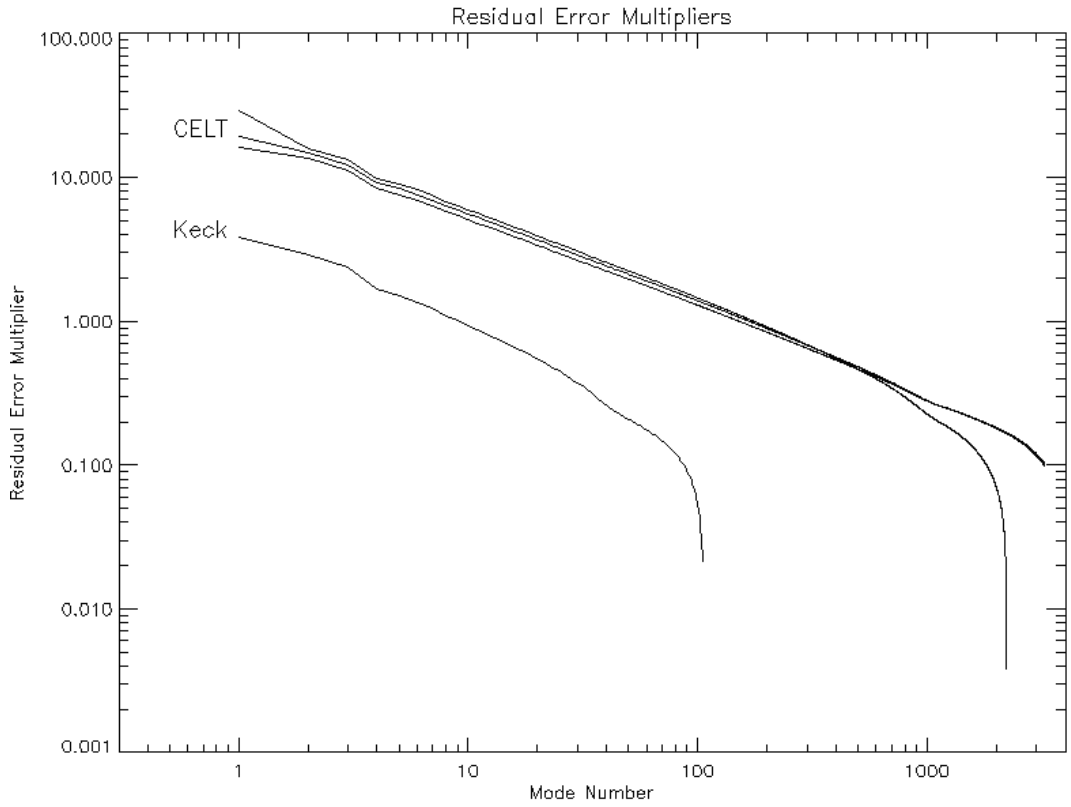


Figure 5-15b. The residual error multiplier after correcting completely the first modes. This multiplier times the rms sensor noise gives the rms actuator value.

Stability of segment phasing is particularly important for AO observations and for infrared observations where single-segment diffraction-limited images are larger than the natural seeing. From the control algorithm Chanan has shown that the rms edge discontinuity is of order unity times the sensor noise, and that the rms adjacent piston errors are also order unity times the sensor noise.

5.4.3 Actuators

Requirements

A conceptual design of the mirror cell exists, and the zenith -65° actuator motions are as large as 2.25 mm (see Section 7.2.4). The shape of the deformation will be dominated by quadratic changes. This implies a maximum actuator velocity of about 140 nm/s.

The gravity-induced motion will be extremely smooth over the primary. The resulting geometric optics image blur induced is under 0.015 arcsec/second. Hence corrections may be applied as slowly as once every 5 seconds and still maintain the seeing-limited image quality of the telescope.

Again, since the changes are very smooth, these errors do not need to be corrected frequently for adaptive optics use. The AO system will easily accommodate more than 10 seconds of deformation. However, for AO, smoothness is the critical concern. The applied motions must be made smoothly, as step errors between segments will degrade the Strehl ratio, and should be kept below 10 nm.

We set a goal that differential actuator errors will be less than 5 nm rms (driven largely by AO, but also to a lesser extent by seeing-limited requirements). Geometric image blur from actuator noise is given by

$$\theta_{rms} = 2.31\ \sigma_a/t \qquad (5\text{-}38)$$

For σ_a = 5 nm, t = 255 mm, θ_{rms} = 0.009 arcsec.

We can get a feeling for necessary update rates with a simple example. Assume actuators have a systematic fractional scale error of β. If the error, δp, in the actuator motion is proportional to the commanded move, p; actuator velocity, V, is 140 nm/second; and the actuator update time is T; then $\delta p = \beta p$ and our noise requirement implies

$$\delta p = \beta TV < 5 \text{ nm rms} \qquad (5\text{-}39)$$

If moves are made to 3%, then T can be as long as ~ 1 second where T is the period of the measurement of the actuator move. This time may be that of the primary mirror control loop, or if individual actuators are under closed-loop control, then it could be that of the individual loop. We expect the primary mirror control loop will update the actuator positions at periods of about 1 second.

The requirements for the actuators used in the Keck telescopes (Mast 1985) are listed in Table 5-11, along with the estimated CELT requirements.

For CELT our baseline requirements are more stringent for noise and resolution than for Keck. Due to the greater gravity-induced deformation of the mirror cell, the required range for the CELT actuators is larger than for Keck.

Table 5-11. Actuator requirements

	Keck	**CELT**
Range	> 0.6 mm	> 2.4 mm
Rms position error (0 Hz)	< 20 nm	< 7 nm
Slew rate	> 10 μm / sec	> 10 μm / sec
Transverse load capacity	> 14 kg	> 5 kg
Axial load capacity	> 150 kg	> 30 kg
Axial stiffness	> 59 N/micron	> 10 N/micron (~100Hz resonance)
Transverse stiffness	> 1.27 N/micron	> 0.1 N/micron
Local power dissipation	< 10 W	< 2 W
Lifetime		$>10^3$ yrs (one failure/3 months)
Survival temperature		-18 to +22 °C
Operating temperature		-6 to +10 °C
Operating humidity		1 to 100% condensing

The Keck actuators carry the axial load of the segments. The CELT baseline segments weigh 74 kg, and the Keck segments weigh 400 kg. Thus, the load requirements are reduced by a factor of 5.4.

We began the study of actuators for CELT with an in-depth and extensive study of design options and suppliers of actuators that could potentially meet our requirements. This study was made by Alan Schier (2001). He investigated actuators based on a broad range of driver technologies: screws (sliding, ball, and roller), piezo-electric, ferro-magnetic, thermal, voice coil; and a variety of motion reduction technologies: Hatheway flexure, hydraulic, levers. After obtaining information from actuator manufacturers with commercial devices and from many people with proposed designs, Schier narrowed the number of candidates to six that he researched in detail. These were then reviewed by a group of experts on actuation. Schier's recommendation, approved by the review group, was that the voice-coil actuator was the most attractive. The second most attractive option was a conventional screw design coupled to a Hatheway motion reducer. Limited resources for the conceptual design phase led us to further explore only the first option. A voice coil design has many attractive features that include, potentially, a very small number of moving parts and potential for a very long life with minimal maintenance. Given the large number of actuators required and the lack of redundancy in the actuator system, these latter features are extremely important.

Lorell and Aubrun (2001) have designed a voice coil actuator for CELT. The central component is a voice coil motor, similar to the main component in most loudspeakers. It has no inherent mechanical stiffness, and only supports the segment when powered. To be a positioning device the actuator uses an internal position sensor (not to be confused with global control edge sensors). A local feedback control system adjusts the voice coil current to maintain the position sensor reading at the value commanded by the global array active control system. This type of actuator has a number of potential advantages. It can be compact and relatively inexpensive. It can have low weight, few moving parts, large force capabilities, minimal maintenance, and potential for damping mechanical resonances using the active control.

There are two major features that were incorporated into the proposed design: 1) a system to off-load the quasi-static forces to reduce the force required and power dissipated by the voice coil, and 2) a lever system that further reduces the forces required for the voice coil and the off-load system.

The off-load system must respond to the changing force of gravity as the telescope zenith angle changes. Design options include completely passive off-loading using a counterweight, active off-loading using a spring adjusted by a small motor to follow the changing gravity force, or a combination of the two.

For detailed design and fabrication of an engineering-model we selected a concept that included levers to reduce forces and a combination of passive and active off-loading. We selected this initial option to learn as much as possible about implementing all these features. In the preliminary design phase we will study system weight, cost, and reliability to determine whether a purely passive design is adequate for CELT.

The final report by Lorell and Aubrun (2001) contains a detailed analytic treatment of the actuator control, including the effects of the compliance of the whiffletree and support structure. In addition, they have included an extensive description of candidate design options and motivation for selecting among them. The report includes a detailed analytic and quantitative analysis of the expected performance of their design under varying disturbance loads, including a modeled wind load. They detail an estimate of the cost of producing 4000 actuators. Figure 5-16 shows the design.

The voice coil (labeled “MIRROR ACTUATOR”) drives an output shaft inside a 50-mm diameter chimney that passes through a support node of the mirror cell or subcell. Spiral flexures are used to carry transverse loads. A counterweight and flexural levers are used to off-load most of the segment weight. A small trim motor and spring are used to trim the counterweight. Measurement of the position of the output shaft by a high precision capacitive sensor is used as input to the PID voice coil current control. Table 5-12 lists some of the major design parameters.

Lorrell and Aubrun made a detailed cost estimate for 4000 actuators. The costs per actuator for six major cost categories are: mechanical parts ($510), electro-mechanical parts ($580), capacitive sensor ($10), electronics ($295), assembly ($200), test/package/ship ($200), and fees ($400). The total is $2195 per actuator.

Table 5-12. Voice-coil actuator design parameters

mirror mass per actuator	25 kg
actuator mass	5 kg
housing length x diameter	18 cm x 14 cm
neck length x diameter	5 to 50 cm x to 5 cm
lever ratio	1/10
counterweight mass minimum	1 kg
position sensor noise equivalent	3.6 nm
actuator stroke	1.2 mm
actuator power constant	5 N / W $^{1/2}$
actuator bandwidth	100 Hz
Off loader	
spring constant	490 N / m
gear ratio	2400:1
shaft radius	10 mm
bandwidth	0.1 Hz
sensor dynamic range	1.2 mm/4 nm = 300,000:1
power dissipated (in power supply)	
under assumed wind loads	150 mW
due to counterweight imbalance	4.2 W / N
power dissipated (in actuator assembly)	1 W

Lorell and Aubrun (2001) have built an engineering-model actuator and delivered it for testing in February 2002. During the next design phase we will test the engineering-model under a variety of load conditions for precision, reliability, and lifetime. The results will be used to evaluate technical risks and as a basis for design changes. During this phase we will also address the schedule and budget risks.

5.4.4 Global Electronics

The global control loop is an over-sampled position loop with 3240 individual nested local loops in the actuators. As described above, the actuator uses an internal sensor to measure the relative actuator position. The local servo loop strives to null the difference between the error value provided by the global loop and the position measured by the displacement sensor.

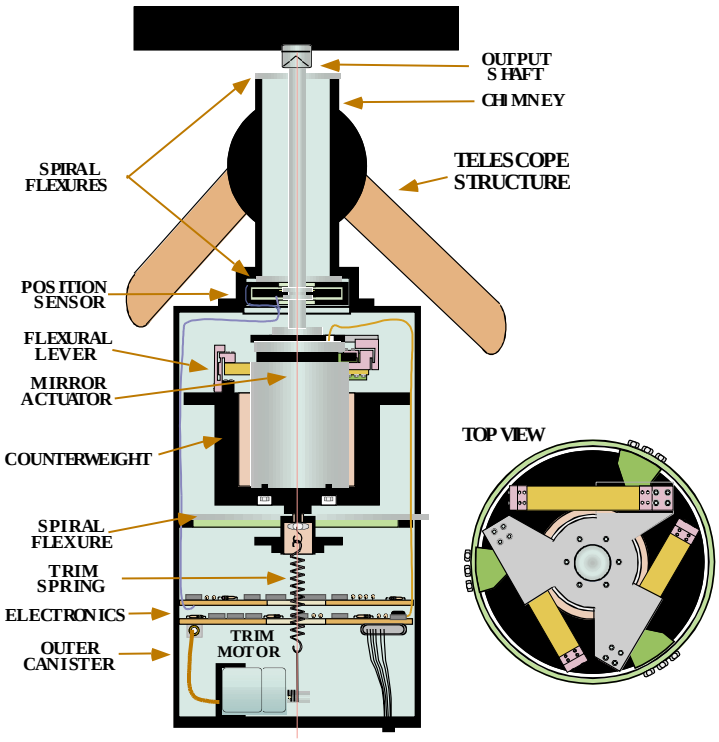


Figure 5-16. Design of voice-coil actuator.

Control system redundancy and system diagnostics

The control system logic and basic performance are described in Sections 5.4.1 and 5.4.2. Since the active control system has 6204 sensors and only 3240 actuators, there is redundancy in the control system. Since the rigid body motion of the primary is not measured by the sensors, there are 2961 extra measurements made by the sensors beyond those strictly needed for control of the primary. These extra measurements provide substantial capability for monitoring the health of the control system, and for redundancy if sensors fail.

For each control cycle, one can calculate the contribution to chi-square from each of the sensors. If the sensors are working properly, then independent of the amount of segment motion desired, each sensor should have approximately the same residual. By simply monitoring these residuals for each sensor, one can immediately determine if a sensor is drifting relative to its calibration, or if it is failing in any other way. Having recognized such failure, one can then note that the sensor electronics or mechanism needs replacement (to be done at a convenient time) and proceed with the control system, ignoring that defective sensor. Deleting a sensor requires recalculating the B matrix, and this is readily done.

The actuators can be monitored in a related fashion. By noting the calculated motion expected of each actuator one can study the smoothness and accuracy of each actuator. The externally induced segment motions are not known for each sample, but they will be smoothly varying, and similar for each segment. The signature of an improperly functioning actuator will be the consistent need for larger corrections than nearby actuators.

For both sensors and actuators, it is thus relatively straightforward to monitor their correct functioning or quantify their degree of malfunctioning. When such malfunctioning becomes excessive the components should be replaced. Whether the actual source of error lies in a mechanical or electrical subcomponent is not determined by this monitoring scheme.

Global loop stability and performance

The global loop stability and performance are governed by three characteristics: 1) the loop gain, which is set by a coefficient in the software, 2) the system delay (i.e., calculation times and filter delay), which need to be small compared to the sample time, and 3) the corrective move, which should be made before the next sample is taken.

System delay comes from processing time, communication time, and delay time. These times need to be reduced to a minimum in order to achieve the best global loop performance. Parallel processing can reduce processing time.

Structural resonances with frequencies that fall in the upper end of the bandpass may be insufficiently controlled to meet the rms error goals. In this case filters might be used to remove these resonances from the sensor signals. The filters will also track the resonance frequency if it changes as the telescope changes elevation angle. Using high Q notch filters or anti-correlation filters will minimize the delay due to this filtering. Both methods track the resonance frequency by periodically making an FFT of the sensor output to identify resonance spikes that exceed a threshold amplitude above the noise. The notch filter algorithm uses the measured resonance frequency and Q. The anti-correlation filter requires both amplitude and phase information from the FFT. Since actuator motion will stimulate the mechanical resonances, they will be coherent with respect to the start of actuator motion. The anti-correlation algorithm will remove the resonant portion of the signal. The anti-correlation filter introduces less delay than the notch filter. In addition to the resonance filtering, a hardware anti-aliasing filter is applied to each sensor signal. Since the anti-aliasing filter will introduce a significant delay below the Nyquist frequency, it is important that this frequency be high. We thus plan in the full system to sample at about 100 times/second, allowing filtering below the Nyquist frequency, and providing enough bandwidth with minimal phase shift to include and possibly affect the lowest mirror cell resonance.

System flexibility

The global loop is a distributed intelligence system with intercommunication links between special function devices. This approach provides the means to change the loop behavior without hardware changes after the telescope is assembled. As an understanding of the detailed response of the structure and environment is acquired, that information can be incorporated into the relevant software modules of the loop.

Diagnostics

The programmability of the various functions of the control hardware permits the system to run self-diagnostic software. The system self-diagnostic software will be run in real time during normal loop functioning, in idle CPU time, and during normal maintenance procedures. Control and monitor microprocessors are assigned to sensor and actuator electronics and to the signal conditioning and array processor modules. By running set data patterns, the software will pinpoint malfunctions to a specific board, which can be rapidly replaced.

Maintenance
A minimum of electronics is placed in the mirror cell. The majority of the electronics is on the Nasmyth platforms or off the telescope. This will provide ease of access to any malfunctioning components. In combination with the diagnostics, the ease of access will minimize loss of observing time when electronic components fail.

Failures
The high multiplicity of components allows and demands a particular program for dealing with component lifetime and failures. All components will be put through an initial burn-in process in order to eliminate infant mortality problems. The multiplicity of components will provide identical (and burned-in) replacements should a failure occur. In the detailed design and implementation of the electronics the lifetime of components will be of the highest priority in component selection. Reducing the part count is the first goal in order to improve the reliability; thermal control will also be essential.

Thermal control
An advantage of having the majority of the electronics on the Nasmyth platforms or in a separate room is the ability to keep a low, constant temperature in the electronics racks. A low chip temperature will be maintained by good component layout and by providing adequate cold airflow between each board. Although there is a high component count, good low-temperature thermal control of active component junction temperatures can give several orders of magnitude improvement in device life.

Ground loops and noise
Widely distributing the electronics would normally make them susceptible to ground loop and induced noise problems. Fiber optic links and a single point ground system will be used to avoid these problems. Power will be distributed in a star configuration, with local isolation and regulation in each electronics rack. Inter-location communication will be over high-bandwidth serial optic links.

Computer room
The global loop algorithm and timing control will be performed here. A signal conditioner will apply the desired offsets to the digitized sensor readings and periodically make a Fourier transform of the sensor outputs to seek the mechanical resonances in the system, and position adaptive filters to remove them from the sensor data as needed. An array processor will make the global control matrix multiplications and also condition the signals to the actuators. Both the signal conditioner and array processor will be monitored and controlled by separate microprocessors.

5.5 Assembly

The assembly of the Keck segment array was made using a pair of high precision theodolites to manually survey coordinate-defining targets on the telescope structure and targets at the locations of the segment support actuators. The survey data was used to calculate errors in the position of a subcell holding three actuators per segment. Manually adjusting the subcell in six degrees of freedom in the mirror cell, and then re-measuring with the theodolites corrected these errors. The final positioning accuracy was limited by the theodolite measurement error of about 100 μm.

A single independent cross check on the absolute accuracy of this system was made using commercial photogrammetry measurements of a variety of targets on the telescope structure and mirror with an rms error of about 30 μm (Dubois 1990).

For CELT, manual measurements will not be feasible. Possible options include

- purchase a photogrammetry system (camera and associated software) for in-house repeated use, or
- purchase an automatic computer-driven system of theodolites that both finds and accurately centers on-targets of approximately known position.

Either of these options will be expensive and require a significant level of in-house expertise to operate effectively. These and other possible options need to be researched in the next design phase.

5.6 Segment Alignment

For reasons of accuracy, reliability, and economy, the determination of the CELT desired sensor readings will be accomplished by an optical, on-sky technique. A fundamental choice has to be made between what we refer to as passive and active methods. In a passive method one reconstructs segment tip, tilt, and/or piston errors (and potentially higher order aberrations) from an image of a star. This is generally done with out-of-focus images, without the benefit of auxiliary optics to facilitate the wavefront error analysis. Curvature sensing (and its physical optics generalization and phase discontinuity sensing) and the Gerchberg Saxton reconstruction are examples of such methods. In an active method, an auxiliary optic is interposed at the pupil or focal plane in order to break the stellar image up into subimages, sorting the photons according to pupil location (as in Shack-Hartmann sensing) or slope error (as in pyramid sensing). Although passive techniques offer the potentially great advantage of requiring no dedicated hardware (the out-of-focus images can be taken with an existing science camera), the CELT baseline is to use a dedicated Shack-Hartmann camera, with a design based upon the Keck phasing camera system (PCS) for sensor calibration. There are several reasons for this choice:

1. The PCS approach has worked successfully for Keck. The existing PCS hardware provides a good starting point design for a CELT phasing camera. The existing PCS software reflects over ten programmer years of development that can be directly applied to the CELT effort.

2. The PCS/Shack-Hartmann algorithms are fundamentally parallel in nature and can therefore in principle accommodate the 30-fold increase in the number of segments of CELT versus Keck. Although the array of microprisms used at Keck would probably be difficult to expand by this factor, we have shown in a series of laboratory measurements that commercially available microlenslet arrays are now of sufficient quality to constitute a viable substitute for microprisms. Other scaling issues (available area on the detector and computation time) are shown below to be manageable.

3. Passive techniques are much more vulnerable to systematic effects associated with segment aberrations than are the Shack-Hartmann and related techniques. An attempt to implement curvature-based segment tip/tilt alignment at Keck failed for this reason.

4. No other alignment technique to date has successfully equaled or even come close to the large (and necessary) segment piston capture range (+/- 30 μm or more) of the PCS broadband phasing algorithm.

We can estimate the required CCD size for Shack-Hartmann phasing of CELT as follows. At Keck the Shack-Hartmann image scale is 6.77 pixels/arcsec, and each subarray corresponding to the diffraction pattern from an inter-segment edge is 33 pixels or 4.88 arcsec on a side. Nearest neighbor patterns are about 60 pixels or 9 arcsec apart. These numbers are not optimal for phasing a very large telescope, where the CCD would be much more crowded with images. The Keck diffraction patterns actually spill over the boundaries of the subarrays to some extent; nearest neighbors should probably be at least 6 arcsec apart. Simulations suggest that the Keck image scale is overly generous, and a scale of 4 pixels per arcsec would probably suffice. For the 1080-segment CELT there are about 70 nearest neighbor diffraction patterns across the full-aperture phasing image. At the above image scale and nearest neighbor separation, this corresponds to 1680 pixels. Thus allowing for appropriate margins, a 2048 x 2048 detector should be adequate.

Probably the most demanding computation associated with the determination of the CELT desired sensor readings would be the inversion of the 1000 (segment) by about 3000 (edge measurements) phasing matrix. We have generated this matrix and inverted it via an SVD algorithm in about 15 minutes using a 670 MHz Alpha 21264 processor. Modest improvements in the algorithm, together with expected gains in computation speed over the next 10 years, should render this computation time negligible.

5.7 Maintenance

Maintenance of the Keck observatory is more costly and complex than was anticipated in the design and construction phases. Informed by the Keck experience we are emphasizing in all designs the goals of efficient and low-cost maintenance. Operations requiring manual intervention at Keck will be carried out remotely at CELT under computer control. Reducing the part count, reducing the number of moving parts, and using long-life components will be given highest priority for all designs. In the next design phase we will begin to write maintenance procedures, failure modes analyses, failure detection programs, and repair requirements for all subsystems. These will be important components in the design of every subsystem.

REFERENCES

Burge, J. 2001. Private communication.

Cabak, J. 2001.“Finite-Element Analysis of a Planetary Stressed Mirror Polishing Fixture Concept.” CELT Technical Note No. 8.

Chanan, G., J. Nelson, T. Mast, P. Wizinowich, and B. Schaefer. 1994a. “The W.M. Keck Telescope Phasing Camera System.” Instrumentation in Astronomy VIII **2198**.

Chanan, G., T. Mast, J. Nelson, R. Cohen, and P. Wizinowich.1994b. “Phasing the Mirror Segments of the W. M. Keck Telescope.” SPIE Proceedings “Advanced Technology Optical Telescope V” **2199**.

Chanan, G., M. Troy, F.G. Dekens, S. Michaels, J. Nelson, T. Mast, and D. Kirkman. 1998. “Phasing the Mirror Segments of the Keck Telescopes: The Broadband Phasing Algorithm.” *Applied Optics* **37**,140-155.

Chanan, G., M. Troy, and E. Sirko. 1999. “Phase discontinuity sensing: a method for phasing segmented mirrors in the infrared.” *Applied Optics* **38**, 704.

Chanan, G., C. Ohara, and M. Troy. 2000. “Phasing the Mirror Segments of the Keck Telescopes II: the Narrowband Phasing Algorithm.” *Applied Optics* **39**, 4706-4714.

Cohen, R. , S. Andreae, A. Biocca, R. Jared, J. Llacer, J. Meng, R. Minor, and M. Orayani. 1990. “The W.M. Keck telescope segmented primary mirror active control system software.” Keck Observatory Report No. 191. SPIE Proceedings, “Advanced Technology Optical Telescopes IV” **1236**.

Dubois, A., W. Irace, T. Mast, E. Romana, B. Schaefer. 1990. “ Installation of the Subcells.” Keck Observatory Report No. 196.

Gunnels, S. 2001. “Concept Design Report Mirror Segment Support System California Extremely Large Telescope.” CELT Report No. 16.

Jared, J., A. Arthur, S. Andreae, A. Biocca, R. Cohen, J. Fuertes, J. Franck, G. Gabor, J. Llacer, T. Mast, J. Meng, T. Merrick, R. Minor, J. Nelson, M. Orayani, P. Salz, B. Schaefer, and C. Witebsky. 1990. “The W.M. Keck Telescope segmented primary mirror active control system.” Keck Observatory Report No. 189. SPIE Proceedings, “Advanced Technology Optical Telescopes IV” **1236** .

Laiterman, L. 2001. “Concept Design for a Planetary Stressed Mirror Polishing Fixture.” CELT Technical Note No. 7.

Lorell, K., and J. Aubrun. 2001. “CELT Primary Mirror Actuators: Phase I Study Report.” CELT Report No. 18.

Lubliner, J., and J. Nelson. 1979. “Stressed Mirror Polishing: A Technique for Producing Non-axisymmetric Mirrors” Keck Observatory Report No. 21 (November 1979); *Applied Optics* **19**, 2332.

Mast, T. 1985. “Actuator Performance Requirements.” Keck Observatory Technical Note No. 125.

Mast, T. 1989. “A User’s Manual for Warping Harnesses.” Keck Observatory Report No. 181.

Mast, T., and J. Nelson. 1986. “Segment Motions and Image Blur Due to Cell Deformations for Various Cell Designs.” Keck Observatory Technical Note No. 177.

Mast, T., and J. Nelson. 2000. “Segmented Mirror Control System Hardware for CELT.” CELT Report No. 6 and SPIE Proceedings “Advanced Technology Optical Telescopes” **4003**.

Mast, T., J. Nelson, and G. Sommargren. 2000. “Primary Mirror Segment Fabrication for CELT.” CELT Report No 5; SPIE Proceedings **4003**.

Mast, T. and J. Nelson. 2001. “Warping Harnesses for CELT.” CELT Technical Note No. 6.

Nelson, Gabor, Hunt, Lubliner, and Mast. 1979. “Stressed Mirror Polishing: Fabrication of an Off-Axis Section of a Paraboloid.” *Applied Optics* **19**, 2341; Keck Observatory Report No. 22.

Nelson, J., and M. Temple-Raston. 1982. “Off-axis Expansions of Conic Surfaces.” Keck Observatory Report No. 91.

Nelson, J. 2001. “ Segment Positioning Actuators - Requirements.” CELT Technical Note No. 5.

Panoskaltsis, V., K. Papoulia, T. Mast, and J. Nelson.1987. “Segment Figure Correction Using Leaf Springs on Whiffletrees.” Keck Observatory Report No. 163.

Schier , Alan. 2001. “Summary of the CELT Mirror Segment Actuator Survey.” CELT Report No. 15.

Sommargren, G., D. Phillion, L. Seppala, and S. Lerner. 2001.“Surface Figure Metrology for CELT Primary Mirror Segments.” CELT Report No. 20.

Zaslawsky, M., and T. Mast. 1989. “Warping Harness Influence Functions.” Keck Observatory Report No. 179.

Chapter 6. Secondary and Tertiary Mirrors

6.1 Introduction

This chapter discusses issues related to the secondary and tertiary mirrors. Because they are much smaller than the primary, the technical challenge is reduced as well. In fact, since 8 m monolith mirrors have been made, and the Southern Astrophysical Research Telescope (SOAR) has a 4.3 m primary, there is abundant recent technology available to produce mirrors of difficulty equal to or greater than those discussed here. However, the convex surface of the secondary mirror does add complexities to its fabrication.

6.2 f/15 Secondary Mirror

The secondary mirror has a diameter set to reflect all light coming from a 20 arcmin field of view. We have elected to use a convex secondary in order to make the tube length as short as is practical (see Table 4-2). Gregorian (concave) secondary mirrors have some advantages, particularly in optical testing, but they lengthen the telescope, and therefore will not be used for CELT. Much of our opinion about the secondary is based on the successful and economical optics for the 4.2 m diameter SOAR telescope.

6.2.1 Optical Fabrication

The secondary mirror is challenging to fabricate for three main reasons: it is convex, it is non-spherical, and the tolerances are tight due to adaptive optics needs.

The convex shape makes optical testing at the center of curvature impractical. Hence a single small test system such as one might use for a concave optic will not work here. We suggest that for the convex optic a system of profilometry and a set of subaperture interferometric tests can be used instead. Profilometry is efficient and should be sufficiently accurate to measure the low spatial frequency surface errors during polishing. The high frequency errors will probably require some form of optical testing. To the extent that the polishing process is axisymmetric, testing along a single radius is enough for guidance during the polishing iterations. Assuming we use a 1 m diameter test optic, only 2-3 optical tests are enough for guidance. Towards the end of the polishing process, more complete testing of the surface is needed, at least a few times. This may require measuring of order 20 subapertures and stitching the data together in order to measure the entire secondary with sufficient accuracy.

Polishing difficulty can be estimated by noting the size of the asphericity of the secondary mirror. If it was spherical it would be much simpler to polish, as a single polishing tool would fit the entire surface. We consider three metrics of difficulty here: the magnitude of the asphericity, the asphericity developed per meter, and the error in the fit of a small spherical polishing tool at the edge of the secondary. For a small tool, the mismatch will be dominated by local astigmatism, and will grow quadratically with tool size. Since Keck Observatory has successfully polished 1.4m diameter secondaries, we will compare these two. (We also include the SOAR primary, but for reference only, as it has different issues due to being concave.)

Secondary	**Aspheric departure**	**Departure/size**	**Astigmatism on 10% tool**
Keck	0.46 mm	6.7e-4	8.8 μm
CELT	1.54 mm	7.8e-4	28.9 μm
SOAR (primary)	0.98 mm	4.7e-4	18.7 μm

We see from these numbers that the CELT secondary is several times more difficult to polish than the Keck secondary and modestly more aspheric than the SOAR primary. One approach to solving this is

to use tools that are a factor of two smaller than those used on Keck (size/mirror diameter). This will increase the time of each polishing run by a factor of a few. Experience with Keck was that testing and analysis time was far greater than the actual polishing time.

The third challenge is related to adaptive optics. On Keck the secondary was made to be good enough for seeing-limited observations. For CELT, the secondary will also be used for adaptive optics, which is both a curse and a blessing. Because the AO system can correct low- to medium-frequency errors, the low spatial frequency errors in the CELT secondary will probably be more loosely constrained than they were for Keck. On the other hand, high frequency errors must be small enough not to degrade the diffraction limit of the telescope. Hence, assuming the pupil AO actuator spacing is about 30 cm, or about 4 cm on the secondary, polishing errors on this scale must be held to ~ 10 nm rms.

Developing a method for fabricating the secondary will be a major activity for the next phase of work.

6.2.2 Passive Mirror Support System

The convex secondary mirror needs to be adequately supported against gravity and other slowly varying loads. Although the details of the mirror are not yet defined, we make some simple assumptions that lead us to a representative design. From Chapter 4 (Optical Design) we see that the desired clear diameter for the secondary mirror is 3.96 m. We assume that the mirror material is 0.10 m thick, select Zerodur as the candidate material, and assume a meniscus shape. The sagitta of the secondary is 0.158 m, a bit more than our assumed thickness. The glass mass is about 3.2 tons.

The axial support, with 60 optimized supports, will have a gravity sag of about 15 nm rms surface (Nelson, et al., 1982). Since the primary mirror maps onto a 3.73-meter diameter region of the secondary, the secondary support spacing is equivalent to a spacing of about 4.1 m on the primary. This is large compared to any proposed AO DM actuator spacing, so from a perspective of adaptive optics, this 15 nm rms surface error will be readily reduced by any AO system. Similarly, the slope errors of this secondary will be smaller than the slope errors on the primary mirror segments and will have very little effect on the seeing-limited image quality. Typical image diameters can be estimated as about 8 times the surface height variations/support spacing (as projected onto the primary), or in this case, 0.006 arcsec. This is about 20% of the allowed error defined in Chapter 11 (Expected Image and Wavefront Quality). We conclude that for a mirror thickness of 0.10 m with 60 axial supports, the gravity driven deformations are completely acceptable.

However, there is a subtlety we must consider for AO. With multi-conjugate AO (MCAO) systems (See Chapter 9, Adaptive Optics), we must consider the effective height of the aberrations. The secondary mirror appears to be about 320 meters below the primary when seen from a star through the primary; and therefore its aberrations do not resemble any atmospheric aberrations. However, this distance is relatively small compared to the scale height of aberrations in the atmosphere, so correcting it as though it came from the primary should be adequate.

Our concept for the in-plane support is less detailed. However, it is clear from the lateral supports of other large mirrors such as those at the SOAR, Gemini, and VLT observatories, that this is not a particularly difficult problem. Here we assume that a combination of reactive in-plane forces at the 60 axial supports and lateral supports along the perimeter will suffice to limit the deformations to an acceptable level.

Although we do not have a detailed design, we assume the support system will weigh about the same as the glass itself, 3.2 tons. The active system discussed below will probably add 2 tons to the mass.

6.2.3 Active Mirror Support Systems

Active control of mirror shape

The passive support system for the secondary is likely to require significant precision and accuracy of design and construction if it is to work as desired. With less expense, the secondary will likely have some residual low spatial frequency errors that are driven by gravity. There are two approaches that might be used to correct these errors.

To the extent the errors are repeatable, they will be "automatically" included in any calibration lookup table for the gravity offsets for the primary mirror. Since the primary mirror has 1080 segments, their piston, tip, tilt adjustments will do an excellent job of correcting any low frequency errors in the secondary.

An alternative is to include active control in the 60 axial supports for the secondary. These 60 supports can then apply variable axial forces in response to external measurements of the deformations of the secondary, where the force is that needed to restore the secondary to its design shape. These corrections could be based on a lookup table or on real-time measurements. Lookup table correction of the secondary is used in our reference design.

To select the optimal approach will require addressing important control issues during the preliminary design phase. The key problem is how to distinguish between optical errors of the primary and those of the secondary. Simply monitoring a star on-axis measures the optical sum of primary and secondary errors. This will not allow an unambiguous determination of the secondary support errors since the primary is also in principle subject to support errors. The primary mirror will have segment edge sensors to stabilize the shape of the primary. These will provide a stable and smooth mirror at relatively high spatial frequencies, but to the extent that sensor errors will develop, they will largely express themselves as low spatial frequency errors in the primary. To an on-axis wavefront sensor, these will look very similar to secondary mirror support errors.

In order to resolve this ambiguity, we will need to include wavefront errors from a few off-axis stars in addition to the on-axis wavefront errors. This should work in principle, because the off-axis stars share the same primary, but have their optical footprints displaced on the secondary by up to 0.15 meters when the star is 10 arcmin off-axis. Even with this, the lowest frequency error sources may be difficult to separate between primary and secondary.

It is important to recognize that to the extent it is difficult to properly separate the sources of error, it is also not important that we do so correctly. The sum of the two mirror errors must be measured and corrected, but the division between the two mirrors must be understood only to suitable, not necessarily exact, accuracy.

Active control of mirror position

A second class of active support will be needed: active control of the rigid body motions of the secondary. We expect from the preliminary structural analysis of the telescope structure that the secondary mirror will move up to 20 mm with respect to the optical axis of the primary (See Chapter 7, Telescope Structure).

Given the need to actively control motions up to 20 mm, we are considering a hexapod support for the secondary mirror and its passive/active support. As we learned from Chapter 4, the secondary mirror needs to be correctly positioned with respect to the primary optical axis to an accuracy of roughly 0.1 mm. Lateral motions of 100 μm produce 0.4 arcsec of image motion and 0.04 arcsec of coma. In addition, any errors should develop smoothly, down to a scale of roughly 50 μm. It is also convenient to use the secondary to significantly change the focus of the telescope. An axial motion of the secondary by 1.0 mm will produce a defocus image diameter of 3 arcsec. Thermal changes of secondary position are likely to be less than 5 mm for a 10°C change of temperature. Given all the above, we expect that 100 mm of motion range is desirable. Actuators with a range of 100 mm and a smoothness/step size of 50 μm (2000:1) should be readily available.

As discussed in Section 4.4, the aberrations from misalignment will be image motion defocus and coma. Again, there is ambiguity as to where such measured errors should be corrected. Image motion can result from moving the telescope, moving the secondary, and moving the tertiary. Image coma can be altered by changing the shape of the primary or the secondary or by moving the secondary. These ambiguities will be addressed in the preliminary design phase.

Active compensation for wind loads

In addition to gravitational and thermal disturbances, we expect that resistance to wind loading will define important requirements. Even with a dome to shield the telescope, when it is windy and the dome is pointing in unfavorable directions, the wind will buffet the telescope and potentially induce undesirable image motion.

The power spectrum of the wind loads on the telescope is not yet known in any detail. The wind itself will have variability. In addition, when the wind interacts with the enclosure and the slit opening, additional turbulence will be induced. Finally, the wind will blow on the upper parts of the telescope and set up Karman vortices, producing oscillatory lateral forces on the telescope.

We expect to operate the telescope in external winds up to 15 m/s. Wind tunnel tests of the Keck dome indicated that the enclosure reduces the wind speed by at least a factor of two at the top end of the telescope (Kiceniuk and Potter 1986). More recent measurements of the effect of wind loads on the Gemini telescope (Gemini 2000) also indicate comparable reduction. With 7.5 m/s wind speeds on the telescope, we expect there will be significant mechanical disturbances at the wind crossing frequency of 0.25 Hz. In addition, oscillatory vortices being shed by the secondary structure itself will produce power at $\sim 0.2v/D_{sec}$ or 0.4 Hz.

In order to correct for the effects of wind, we plan to make the active control of the secondary mirror position rapid enough to remove most wind effects. We believe that a control bandwidth of the position of the secondary of 1 Hz is practical, and will attenuate the wind disturbance adequately.

Wind loads are discussed more extensively in Chapter 7 (Telescope Structure) and will be a major subject of study in the next phase of work for CELT.

Chopping and Nodding

Where the sky background is large, such as at thermal infrared wavelengths, small detector sensitivity changes can cause apparent background variations to dominate over valid astronomical signatures. In these circumstances it has been found useful to chop the astronomical image between two or more

locations on the detector. This oscillatory motion effectively removes drifts in sensitivity that occur at frequencies below the chopping frequency. We expect that this technique will be useful for CELT. It is clearly a challenge to move the secondary mirror rapidly. The upper part of the telescope tube is likely to have its lowest natural frequency above 5 Hz. We expect that with some care we will be able to produce approximate square wave motions of the secondary above 1 Hz. Chopping amplitudes of 10 arcsec on the sky will be quite useful. These imply that rapid actuator moves of roughly 200 μm will be needed (10-arcsec image motion). Repeatability of 0.05 arcsec will require actuator repeatability of roughly 1 μm.

The actuator requirement for chopping smoothness (1 μm) and the actuator total stroke (100 mm) imply a dynamic range of 100,000:1, a specification that may be difficult to achieve in a single device. It may be that the secondary actuators will operate with two stages; one coarse, and on top of it a smaller stroke device that is smooth and rapid. This will be investigated during the next phase of work.

6.3 Adaptive Secondary Mirror

We are considering the option of having an adaptive secondary mirror as well as the reference design secondary. Such a secondary would be interchangeable with the reference design secondary. Although we have neither a design nor requirements for such a mirror at this time, it will probably be no larger than 2 m in diameter.

The tolerances for positioning this secondary will be similar to those of the baseline mirror described in Section 6.2.

Since the main task of such an additional secondary will be to apply low order correction to the wavefront (no more than 1000 actuators), its gravity deformations will not be a major concern. The secondary will be controlled by the AO system that will be measuring net wavefront errors from the atmosphere and all other systems, and applying these to the secondary. The key potential virtue of such a secondary is a reduction in net emmissivity.

Design details and desirability will be investigated during the next phase of work.

6.4 Tertiary Mirror

As discussed in Chapter 4, the flat tertiary is roughly 3.1 m x 4.4 m and assumed to be 0.1 m thick. We assume the material is Zerodur. This leads to a glass mass of about 2.7 tons. As before, we assume the passive support will be equal in mass, or 2.7 tons. The active motion mechanisms will probably add 2 tons to the total mass.

6.4.1 Optical Fabrication

Since the tertiary is flat, the optical testing will be relatively simple. The footprint of a single star on the tertiary is 1.33 m and over that region, the tertiary should be flat enough to meet the error budget for seeing limited images.

When AO is used, the requirements are somewhat different. Although one might think the AO system could correct for optical errors in the tertiary, with MCAO and a finite field of view, this is not entirely true. As discussed in the next section, the MCAO system has no DMs conjugate to the tertiary, hence

correction is limited. Over the required 2 arcmin FOV, the star footprint on the tertiary will move up to 0.09 m. Thus the surface shape differences between star footprints displaced by up to 0.09 m should be under about 10 nm. This implies that the tertiary mirror needs to be rather smooth, but AO does not impose unusually strict requirements for low spatial frequency surface errors.

Large polishing tools should work effectively, since a flat is a specific radius sphere; hence, the tertiary should be significantly easier to polish than the secondary mirror.

6.4.2 Passive and Active Support

The tertiary mirror is comparable in size to the secondary, and as a result the passive and active support for this mirror may be similar. Because the mirror is closer to the focal plane, some tolerances will be looser than the secondary. However, others involving the wavefront will be similar.

We note that the footprint diameter of a star on the tertiary is only projected at 1.33 m, much smaller than the star's footprint on the secondary. This implies that the spatial frequency of optical fabrication wavefront errors will tend to be lower than on the secondary.

When we consider correcting such errors with an adaptive optics system, we must be careful about the location of the errors. The secondary appears to be relatively close to the pupil/primary of the optical system, so an MCAO system that corrects for primary mirror errors and local dome seeing is also likely to correct for low spatial frequency errors in the secondary. The tertiary, on the other hand, appears to be about 6700 m below the primary, so its wavefront errors will not be well corrected by the MCAO deformable mirrors. These DMs are likely to be conjugate to the primary, and to heights in the atmosphere of roughly +3km to +10km.

The result of this fact is that the shape of the tertiary needs to be well controlled and not just correctable by the AO system. In particular, wavefront errors occurring over a star's footprint (~ 1.33 m) should be small, at the level of 15 nm rms.

6.4.3 Motion Requirements

We expect the tertiary mirror to be articulated about two axes in order to direct light to Nasmyth instruments not directly on the elevation axis. One rotation axis is the primary optical axis. The range of this rotation must span 180° in order to reach both Nasmyth platforms and slightly more to reach the proposed off-elevation-axis instruments. The second rotation axis is perpendicular to the primary optical axis and corresponds to the y axis when the tertiary is directing light along the elevation (x) axis. The range of this rotation is much less, probably only enough to move the focus 5 m above or below the elevation axis, or roughly ±14°. It is desirable that these two rotation axes intersect at the center of the optical surface of the tertiary. It is probable that rotary bearings will be used for this motion control. Rotary encoders along with simple geared motors should be sufficient for position sensing and motion actuation. Geometrical and mathematical details are given by Kuhlen (2001).

6.5 Primary, Secondary, Tertiary Control

As mentioned in the discussion above of the secondary motion, the motion logic and accuracy requirements are subtle and not yet resolved. In essence the issue is this: Several different actuation mechanisms cause image motion. Is it necessary that all have sufficient smoothness to control the

image position to its allowed tolerances of roughly 0.05 arcsec, or is it enough that only one of these mechanisms has this smoothness of control? Given that we will have a guider, any image errors will be sensed and rapidly corrected by whatever smooth mechanism is chosen for this work. The size and duration of allowable transitory errors is not yet resolved. This issue will be studied in some detail during the next phase of design work.

REFERENCES

Gemini South 8-m Optical Telescope, Final Report. Modal Analysis and Controls Laboratory Report No. -05-08570-001. CD-ROM produced by David R. Smith: "Wind Tests Data from Gemini South." May 13-14, 2000. Mechanical Engineering Department, University of Massachusetts: Lowell, Massachusetts.

Kiceniuk, Taras, and Kent Potter. 1986. "Internal Air Flow Patterns for the Keck 10-Meter Telescope Observatory Dome." Keck Observatory Report No. 166.

Kuhlen, Mike. 2001 "The Tertiary Mirror Equations of Motion of an 'off-altitude-axis' Nasmyth Focus." CELT Report No. 22.

Nelson, Jerry, Jacob Lubliner, Terry Mast.1982. "Telescope Mirror Supports: Plate Deflections on Point Supports." Keck Observatory Report No. 74; SPIE Proceedings **332**, 212, "Advanced Technology Optical Telescopes."

Chapter 7. Telescope Structure

7.1 Introduction

The telescope structure has the following main functions:

- Support the primary, secondary, and tertiary mirrors; the AO systems; and the instruments.
- Point the optical axis to any object in the sky and track it over time.
- Allow access for servicing the optics and instruments.
- Maintain support during operations against the expected changes in temperature, humidity, wind loads.
- Survive or operate in all expected environmental situations.
- Minimize the obscuration of light and the emission of thermal infrared radiation.

At this stage, the telescope structure design is incomplete. The design is based on the evolution of several intelligently selected approaches, but is not the result of an exhaustive investigation. We are currently optimizing the structure design to be consistent with many optical and mechanical constraints. We present here the current status and expect that major changes in the topology and geometry of the structure are likely to occur in the future.

The primary-secondary alignment requirements are roughly 0.1 mm (see Section 4.4.2), and segment support positioning requirements are under 100 nm. Ideally, we would like the structure alone to rigidly support the optics in their correct positions as it changes the direction of the optical axis and as its temperature changes. In practice, this is impossible. The next two sections calculate the rough magnitude of the gravity and temperature effects. We conclude from these calculations that active control of the optics is necessary.

7.1.1 Effects of Gravity Loads

Since gravity loads will vary as the telescope elevation angle changes, self-weight deflection of the structure will be at the very least (self-weight compression of a rod)

$$\delta = \rho g L^2 / (2E) \qquad (7\text{-}1)$$

where δ is the deflection
g is the gravitational acceleration (9.8 m/s^2)
ρ is the material density (for steel $\rho = 7.8 \times 10^3$ kg/m^3)
L is the characteristic size of the structure (~ 50m)
E is the material elastic modulus (steel $E = 19.3 \times 10^{10}$ N/m^2)
assuming the above values gives $\delta = 0.5$ mm. Any practical structure is likely to be 1-2 orders of magnitude worse than this.

For an additional example, consider a simply supported horizontal steel bar 50 m long and 10 m deep. It will deform 3.9 mm under its own weight. Deflections of structures like this grow as the 4th power of the length and inversely as the 2nd power of the depth.

These two static examples indicate the lower limit of the gravity driven effects. Designing a structure with varying gravity direction and with all the constraints of the telescope geometry will produce a structure that is inevitably more compliant than these static examples. A useful reference on the general gravity and thermal deformation limits is von Horner (1967).

7.1.2 Effects of Temperature Changes

Temperature changes will cause dimensional changes of

$$\delta = L\alpha\Delta T \qquad (7\text{-}2)$$

where α is the coefficient of thermal expansion of the material (steel $\alpha = 1.2 \times 10^{-5}$ /°C)
ΔT is the temperature change.

Consider an example where the temperature changes by 10°C. In this circumstance the outermost segments will move radially by $\delta \sim 1.8$ mm. For the optical design of CELT, such motions are optically unimportant.

Temperature gradients also cause dimensional changes. For a single material the dimensional changes will be stress free. For a constant gradient of temperature experienced by a homogeneous material, straight lines normal to the direction of the gradient are deformed into arcs of circles, where the radius of the circle is given by

$$R_T = 1 / (\alpha \nabla T) \qquad (7\text{-}3)$$

As an example, if ∇T is 0.1°C/m, then for steel, $R_T = 833{,}000$ m. If we had such a temperature gradient along the optical axis, then the primary mirror structure would develop a temperature driven sag of $s = R^2 / 2R_T = 0.14$ mm. With active optics, this motion can be eliminated. Such a gradient perpendicular to the optical axis would cause the secondary to decenter by 1 mm. Again, with active optics this motion can be eliminated.

7.1.3 Need for Active Control

Thus we see that both thermal and gravitational disturbances will be sufficiently large that careful balancing of the effects of these disturbances in the structural design and active control of the optical element positions are desirable. Even if the segments are “floated” in some ideal fashion thermal deformations will require active control.

Traditionally, optical telescopes have had mechanical designs that passively balance the motions of the primary and secondary mirrors so they maintained good relative optical alignment. The Serrurier truss, initially developed for the Hale 5-m telescope, accomplished this. More modern telescopes have also implemented this idea. For CELT, the motions are sufficiently large that we believe that balancing will be impractical, and that active positioning of the optics will be required. Further, given the success of active control of optics in other facilities (Keck, Gemini, VLT) we accept that active control is practical and more affordable than fine tuning of the structure itself, and also far more effective. Thus for CELT we have made a key decision. We assume that the structure should be built for maximum stiffness (to minimize wind disturbances and improve motion control) and that balancing of optical motions will be achieved actively.

At Keck the primary mirror segment positions are actively controlled with edge sensors and actuators. This system has worked extremely well, and effectively makes a rigid mirror supported by a steel mirror cell. The actuators work against this stiff structure to hold the mirrors in their desired positions.

7.1.4 Telescope Motion Control

We expect it will take significantly longer to move CELT than it takes to move a smaller telescope such as Keck. The power required to move the telescope a given angular distance in a given time varies as L^5 (assuming the mass grows as L^3). In addition, natural structural frequencies are generally getting smaller as L^{-1}, so dynamical limitations become more important as the telescope grows in size.

7.1.5 Geometrical Trade-offs

Several optical parameters drive the structural design. The diameter of the primary mirror, the primary-secondary distance, the location of the elevation axis, the location of the final focus, and the associated instrument sizes are the most important.

In most previous optical telescopes, the primary mirror and its support cell form one end of what is commonly called the telescope tube, while the secondary mirror is held at the opposite end. A structure around the perimeter of the tube, containing the center of gravity of the tube, then forms the attachment to the rest of the structure, which is commonly called the base, yoke, or alidade. This connection forms a moving joint, and the motion defines the elevation axis. The yoke rotates about a vertical axis, relative to the telescope pier (fixed to the ground). These two rotation axes form the elevation-azimuth motion coordinate system.

For the design of larger telescopes, it is structurally advantageous to use a smaller focal ratio for the primary mirror. This is balanced by an increased difficulty in fabricating and aligning the optical surfaces. The Hale telescope is f/3.3, the Keck telescopes are f/1.75 and the current plan for CELT is f/1.5. In contrast, radio telescopes, with much coarser optical tolerances, are less difficult to fabricate and align, and use primary mirrors at f/0.4.

For shorter focal ratios, there are two design options worth considering. First, it becomes practical to consider placing the elevation axis behind the primary, rather than in front of it. When the light beam goes to a Nasmyth platform it is usually required that the elevation axis be significantly behind the primary, so the light beam outward along the elevation axis does not interfere with the support structure. Second, supporting the primary mirror directly behind the surface becomes more efficient than holding the tube only at its perimeter.

Radio telescopes typically place the elevation axis behind the primary, using counterweights to move the center-of-gravity to the elevation axis. This allows the primary mirror support structure to be efficiently developed directly behind the mirror itself. Because the focal length is so short, the size of an enclosure is driven by the diameter, not the length of the telescope.

For CELT we are in a transitional region. If the primary focal ratio were f/1 or less, it might be attractive to move the elevation axis behind the primary. However, at the moment we believe that optical fabrication and segment positioning difficulties preclude this solution. A study by Meinel and Meinel (2000) explores the possibilities of placing the elevation axis behind the primary.

We plan to directly support the primary and its cell from behind the cell and on the yoke, rather than carry the loads up through the perimeter of the tube and over to the yoke. Thus we plan to have large bearing surfaces directly behind the primary, to both transfer the load to the yoke and to define the elevation axis.

This kind of structure is not amenable to the design optimization methods of a Serrurier truss. We expect that the secondary support structure will be developed directly from the mirror cell and support bearing surfaces, rather than from an intermediate elevation ring.

7.2 Primary Mirror Cell

The primary mirror cell supports the primary mirror segments and carries their loads to the yoke. The cell is also responsible for attaching to the upper tube (the support structure for the secondary mirror) and carrying those loads back to the bearings that allow the motion around the elevation axis.

7.2.1 General Objectives

The mirror cell must adequately support the segments against the deforming influences of gravity and temperature changes. It must conveniently attach to the mirror segment system on one end and to the elevation bearings on the other. Since the cell deformations will be larger than the segment alignment tolerances, active control of the segment positions will be needed. Thus a key requirement of the mirror cell design is that its deformations be smooth and not exceed the motion range of the segment support actuators. The range about the average displacement is the key parameter. Currently we are designing the actuators to have a range of 2 mm. In-plane motions must also be limited since we do not plan to actively control these three degrees of freedom. This system has been very successful at Keck.

The upper parts of the mirror cell must allow accurate positioning of the segments, and also allow easy installation and removal of the segments. In addition, servicing requirements must be met, from periodic cleaning of the optical surfaces to replacement of broken segment support actuators and sensors.

We also want the mirror cell to have the lowest practical mass and to be designed for economy of construction and erection. The low mass will reduce the adverse impact on the thermal environment of the telescope and ease the requirements on the motion control system.

7.2.2 Clusters

In our reference optical design we have 1080 hexagonal mirror segments. Because this is a very large number (Keck has 36 segments, Hobby-Eberly Telescope has 91) we do not think it is practical to periodically install and remove these segments one at a time. As a result we have decided to handle and support the segments in groups, called clusters. We expect a standard cluster will consist of 19 hexagonal segments. Thus we will need approximately 60 clusters to make up the primary mirror. A conceptual drawing of the cluster pattern is shown in Figure 7-1. At the periphery we will use partial clusters.

A simple two-layer truss will support a cluster of 19 segments, and this truss will connect to the main part of the mirror cell at three points. The cluster truss is shown schematically in Figure 7-2. Each cluster truss will weigh roughly 745 kg and support a segment mass of 1560 kg plus segment supports. The function of the cluster is to support the segments and hold them stiffly in the correct position, and be easily installed or removed from the mirror cell below. Accuracy is important here, as inter-segment gaps are only 2 mm. The on-cluster segment positioning must be done to a modest fraction of a mm, and the cluster installation on the mirror cell must be at the same level of accuracy.

A preliminary design of the cluster truss is very stiff. It has a lowest frequency on its three support points of about 30 Hz., with displacements perpendicular to the cluster plane. The maximum gravity-driven deflections from zenith to horizon are about 0.150 mm. Details are given by Medwadowski (2001a).

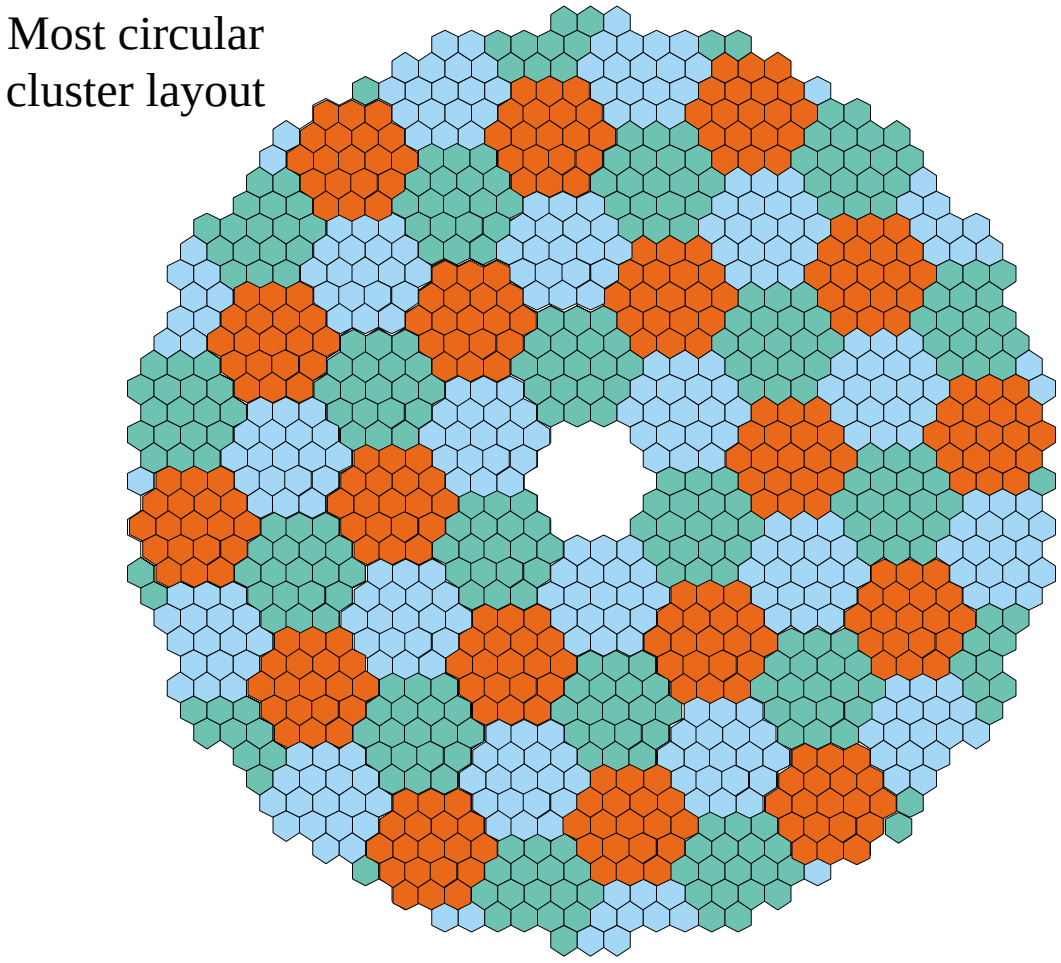


Figure 7-1. Coloring clusters in three different colors indicates the cluster pattern of the primary. Sixty clusters are used, with each cluster (typically) holding 19 segments.

At this time we have not developed the details for assembling, installing, and removing the clusters. This is an important activity for the next design phase.

7.2.3 Cell and Subcell Topology

Each segment will have a subcell with which to interface to the cluster. The subcell will hold the actuators (3/segment) and the lateral support system. The subcell will be a simple and lightweight spaceframe, similar to those used for Keck, and will weigh roughly 8 kg.

Medwadowski (2001a) has studied a subcell geometry. This loaded structure has a lowest natural frequency of 43 Hz.

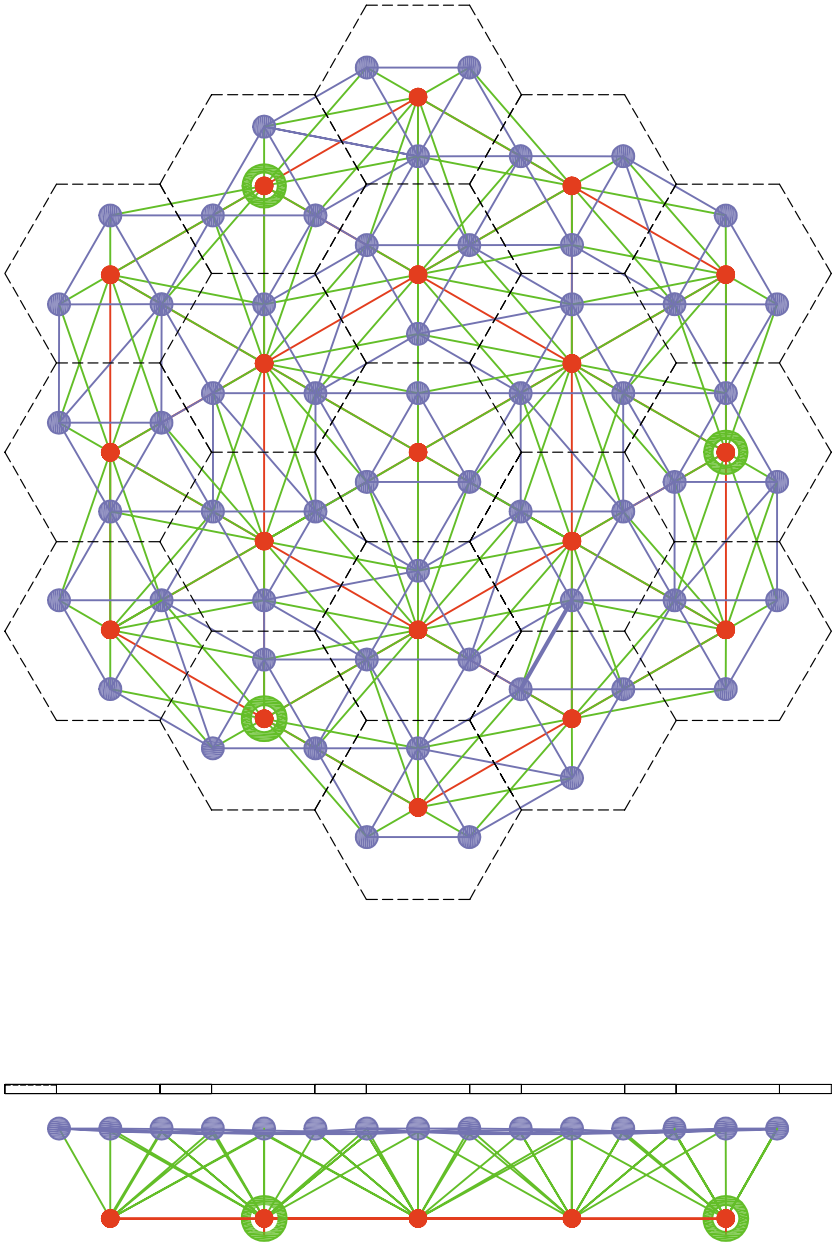


Figure 7-2. The cluster structure is shown in plan and elevation views. The blue nodes represent actuator locations in the upper layer of the truss, and blue lines interconnect them in the upper layer. Red nodes indicate the nodes of the lower layer of the structure, and red lines show the interconnection of these nodes in the lower layer. These connect all nearest neighbor nodes in this layer, although the colors of the connections are not always clear, as green diagonal members occasionally overlap them. Green lines indicate the diagonals between upper and lower layers. Three lower nodes are circles and indicate the connection to the rest of the mirror cell. The segments are indicated for scale. The segment detailed support is omitted.

7.2.4 Cell Dimensions and Performance

The design and performance of the cell and tube are described by Medwadowski (2001b, d). The cell as a whole must stiffly and accurately support the clusters at the top cell surface and must transfer the load onto elevation journals at the bottom. The elevation journals are circular arcs with centers on the elevation axis, 3.5 m in front of the vertex of the primary. A side view of the cell is shown in Figure 7-3. We can see further details of the structure with a front view of the cell, shown in Figure 7-4. The current design has two elevation journals. We have explored a system with four journals that has certain advantages, and in the next phase of work we will continue this exploration.

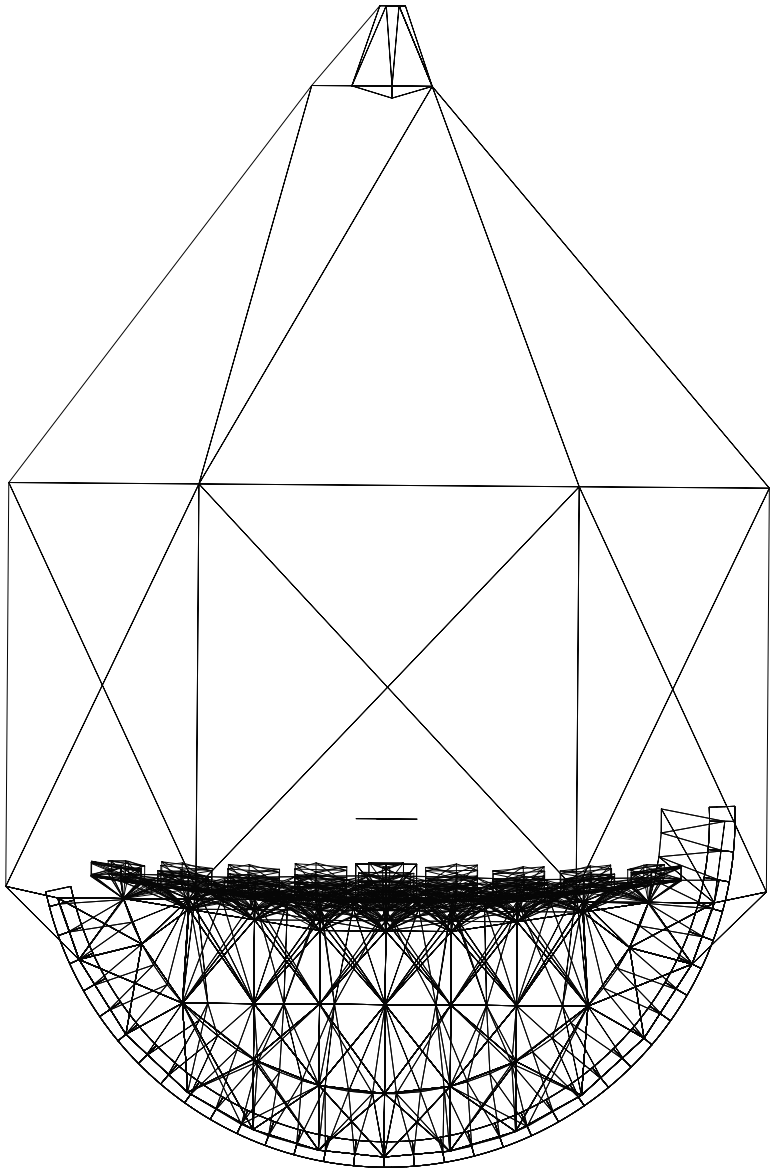


Figure 7-3. A side view of the telescope tube. The elevation journals (circular arcs) are visible, as is the mirror support cell.

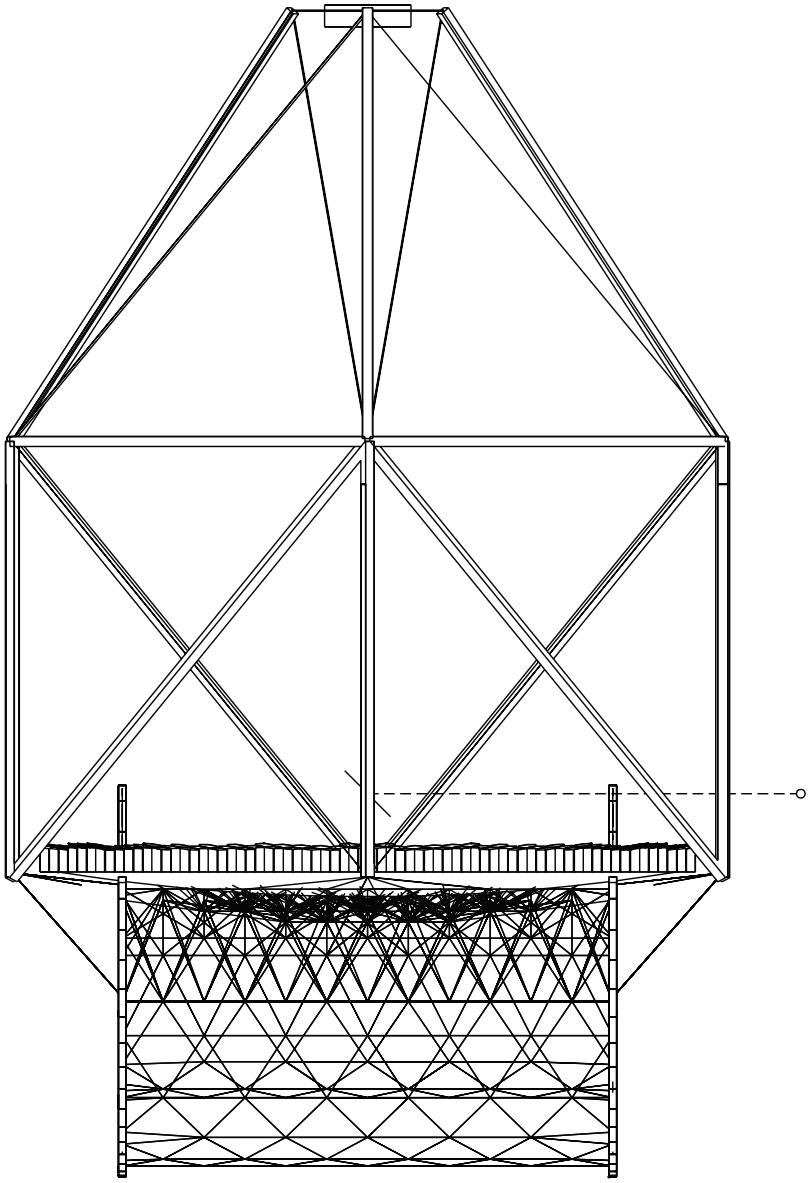


Figure 7-4. A front view of the telescope tube. This shows the mirror cell structure from another perspective.

At the top layer the scale length of the structural elements is 2.5 m (the distance between cluster attachments). As we go deeper into the cell it changes into a pattern compatible with the circular elevation journals. The characteristic structure size now becomes the size of the clusters, about 4 m, and follows a triangular structure. This can be seen in the side view of the structure shown in Figure 7-3.

The critical task of this cell is to hold the segments in the proper location against varying gravity loads. Since the subcell and cluster structure are extremely stiff, we can estimate the performance by looking at the motions of the cluster attachment points.

At the zenith, the average gravity-driven displacement along the optical axis is 2.87 mm. The peak-to-valley range about this average is 13.7 mm. This range by itself is not critical, since the optics can be aligned to the desired position for a single elevation angle. What is critical is the range of motion of each of the attachments over the elevation angle range of the telescope, since this sets the required actuator range. For Keck the corresponding motions are about 1 mm.

As the telescope moves from the zenith to 65 degrees, the primary rises towards the elevation axis by 10.7 mm. By examining deflections at several elevation angles and removing the best fitting plane at each elevation (such rigid body motions are compensated by secondary mirror motions), we can determine the maximum range of any support point. For the present design this is 2.25 mm. This sets the required range of the actuators that support the mirror segments. For Keck this is about 0.6 mm.

In-plane motions are also important. By examining the displacements of the cluster attachments at several elevation angles, we see that the average in-plane motion is largest at 65° zenith angle and is 15.2 mm in the y direction relative to the position at the zenith (z axis is optical axis, x axis is elevation axis). The peak-to-valley deviation from the mean is 3.6 mm. In the coordinate system centered on the average displaced segments, the average segment radial motion is about the same for all elevation angles and is negligible, and the standard deviation of the radial motions is also approximately independent of zenith angle and has a value of 1.0 mm. Radial motions of this size produce very small optical aberrations.

In-plane rotations of the segments about their individual centers are also optically important. For these estimates we have added the cluster rotation (as defined by its three nodes) and the rotation in the translated coordinate system (rotation relative to the center of the primary being the relevant factor). In this coordinate frame the maximum node rotation (an approximation of a segment rotation) is 2×10^{-4} radians, experienced at 65° zenith angle. By rotation we mean the angular change in direction of a line on a segment that originally passed through the center of the primary. This maximum rotation occurs for the outermost attachments.

These motions (displacements and rotations) produce image blur, a consequence of segment in-plane misalignment. The range of motions is well within the error budget for this effect developed in Chapter 11.

At this stage in the design, we are very interested in the in-plane motions, since they give an indication of the practicality of the optical configuration. With the segment misalignments found here, we can consider the possibility of different optical designs. The allowed misalignment varies as f_1^3/a^2. Thus from the perspective of segment misalignment, we could consider segments as large as a = 1.5 m, or f_1 as short as 21 m.

7.2.5 Attachment to Elevation Journals

The stiffness of the elevation journals is critical to achieving adequate performance. The journal surface is attached to the rest of the mirror cell with a truss structure that follows the circular arc of the journal. It should be noted that large journal surfaces introduce structural issues that are significantly different from those of traditional telescope structures. In typical telescope elevation journals, the journal deformations are small and approximately independent of the elevation angle, since the journal is small and circular. Here the journal is only an arc of a circle, and because of its large size, it is a challenge to

make it extremely stiff. This stiffness is likely to dominate the dynamics of the telescope structure as a whole. As mentioned above, the current design has two journals, but we are also exploring designs with more journals.

The elevation journals have a radius of 17.4 m, centered on the elevation axis. They are separated from each other by 22.7 m.

7.3 Secondary Support

The challenge of supporting the secondary is to achieve maximum stiffness while minimizing the blockage of the primary. We have explored several geometrical configurations. Support structures following the design of Keck were tried, but they were massive and had large cross-sectional areas exposed to the winds near the top of the telescope. Single-layer structures were also explored, from tetrapods to pre-tensioned tripods (Medwadowski 2001b, c).

The design structure with the best overall performance is a two-layered structure. The base layer is a hexagonal cylindrical truss, outside of the optical path. The top layer consists of three compression members (square tubing, 0.46 m on a side) positioned radially over the primary, and six tension members (rectangular rod 0.038 m x 0.152 m) that connect the rim of the lower layer with three points that define the corners of the secondary mirror module. Under theoretical conditions where the mirror cell is assumed to be infinitely stiff, the upper tube structure with the secondary mirror system (assumed 10 tons) has a lowest natural frequency of almost 10 Hz. This upper tube structure is shown in side view in Figure 7-3 and 7-4. A detailed description is given by Medwadowski (2001d). Figure 7-5 shows the blockage of the primary by the upper tube.

7.4 Yoke

The yoke supports the telescope tube. The design and performance of the yoke is described by Medwadowski (2002a). The attachment between the tube and the yoke is through hydrostatic bearing pads, where the pads are part of the yoke and the journals are circular arcs forming the lower part of the mirror cell. The yoke must also rotate about a vertical axis, and four azimuth hydrostatic pads allow this motion and carry the telescope load onto the telescope pier. Another essential part of the yoke is the provision for two Nasmyth platforms to carry the scientific instruments. Each of these is a horizontal platform, approximately 15 x 32 m in size. Currently the height of the Nasmyth platform is 3.5 m below the elevation axis, the same as the primary mirror vertex. Each platform has an instrument capacity of 80 tons, and its use is described in Chapter 4. We expect multiple instruments to be located on each Nasmyth platform and each is pointed toward an articulated tertiary mirror. Thus, rotating the tertiary mirror can access multiple instruments without any motion of the scientific instruments.

In the current design there are six bearing pads supporting the tube, three under each journal. This raises an important issue that is not yet completely understood or resolved. By design the pads lie on the surface of a cylinder, centered on the elevation axis. As a rigid body, the cylinder requires six constraints to define its position, three rotations and three displacements. One rotation is the desired rotation about the elevation axis; hence the drive system provides the needed constraint. One displacement is motion along the elevation axis, which we must restrict with additional hydrostatic pads at the faces of the "cylinder." The other four degrees of freedom are constrained by the radial bearing pads. A four-pad system (treated as points here) would properly constrain the other four degrees of freedom (two displacements and two rotations). Additional supports lead to an indeterminate support

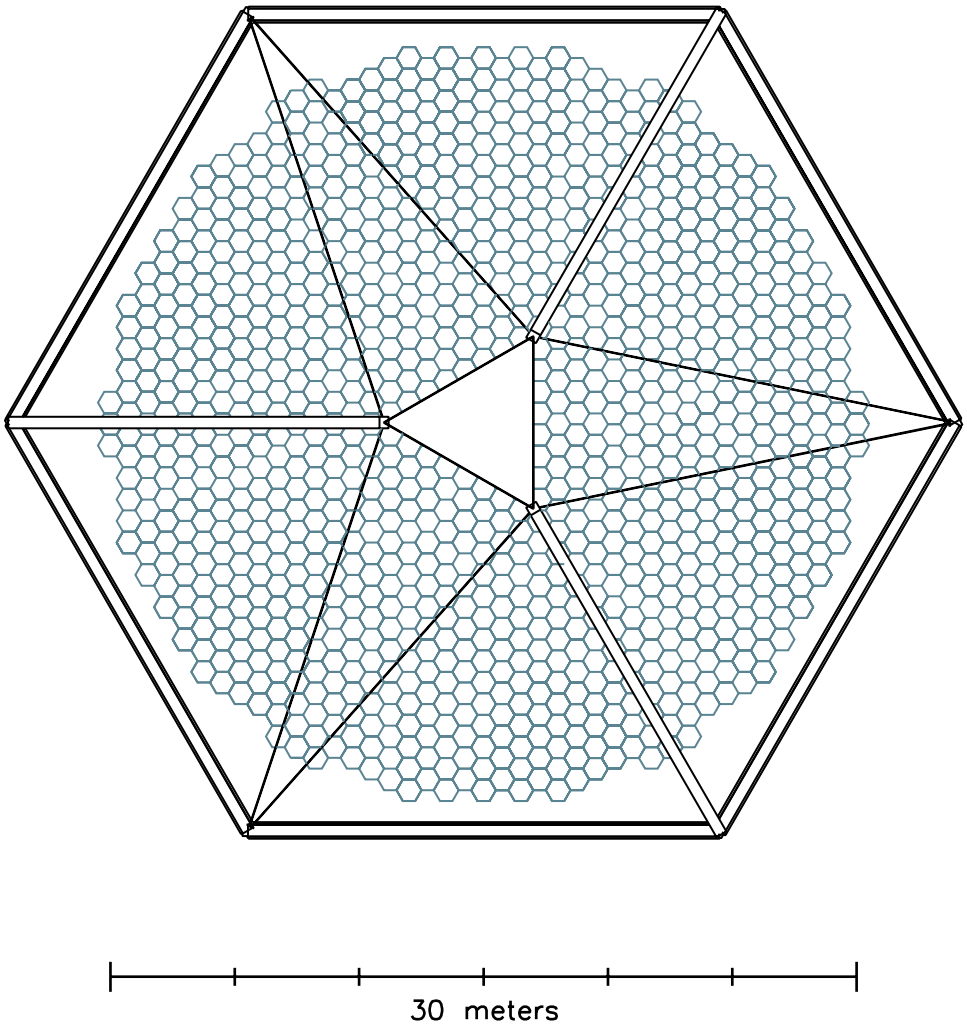


Figure 7-5. View of the primary mirror from a star, showing the blockage caused by the upper tube structure that supports the secondary mirror.

where the loads at the supports are not well defined. However, the tube structure is relatively flexible, thus for small displacement errors it will generally conform to the assumed fixed positions of the pads. Gravity loads on these six pads will be nominally constant with changing telescope zenith angle. However, journal fabrication and pad installation errors or structure temperature differences will cause these six pads to move from the desired cylindrical surface. In this case the tube structure will deform to maintain contact. The potential consequence of this is that the mirror cell will deform and additional actuator range will be needed. This is important because segment support actuators have a limited range that isdifficult to increase. A four-pad system would not suffer from this concern, but a six-pad system does. This is a quantitative issue, but it may also be possible to couple two pads together so one of them can change its height as a function of pressure variations. This would restore the system to a quasi-kinematic design.

A front view of the yoke is shown in Figure 7-6. An end view of the yoke is shown in Figure 7-7. A top view of the yoke is shown in Figure 7-8. The six elevation axis bearings are indicated as small blue circles in these drawings. On each journal, two bearings are symmetrically separated by 50° while the third is placed an additional 20° along the journal. The two journals are separated by 22.7 m along the x axis.

In azimuth the four hydrostatic pads form a rectangle, 23.3m , x 22.7m. The azimuth journal surface is 25.35 m below the elevation axis. The four azimuth pads are indicated as small red circles in Figures 7-6, 7-7 and 7-8.

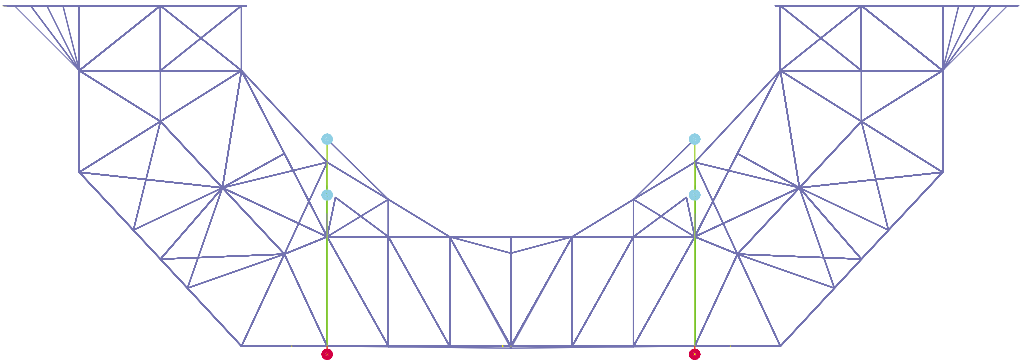


Figure 7-6. Front view of the yoke, showing the structure supporting the Nasmyth platforms and also indicating the elevation bearing pads and the azimuth bearing pads as small circles.

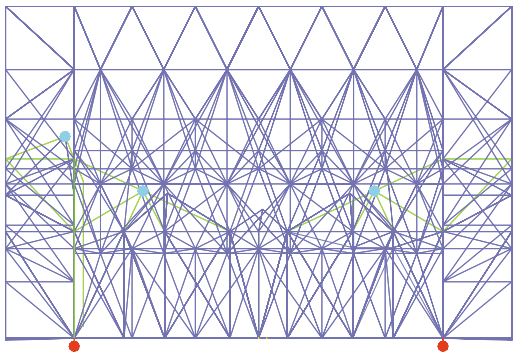


Figure 7-7. End view of the yoke, showing its structure. Again, small circles indicate the elevation and azimuth bearing pads.

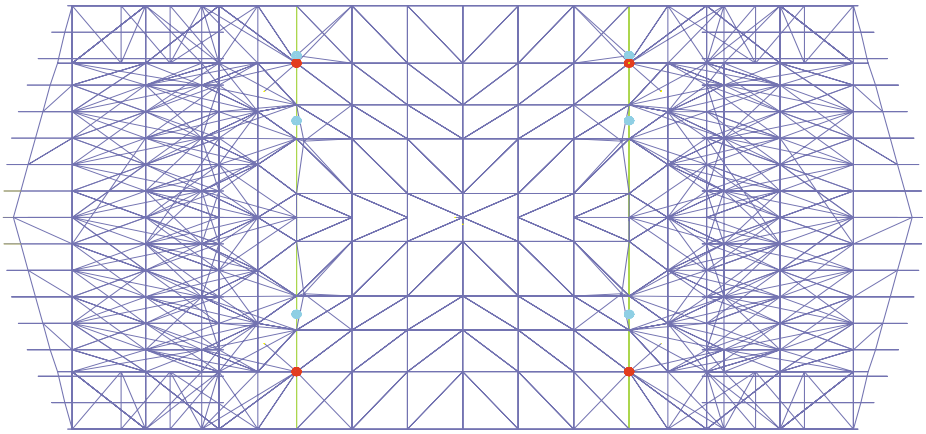


Figure 7-8. Top view of the yoke, showing its structure. Again, small circles indicate the elevation and azimuth bearing pads.

The four azimuth hydrostatic pads are designed to carry the vertical loads and will constrain three degrees of freedom of the yoke (piston, tip, tilt). Once again we are over-constrained, and will need to study the adverse consequences of azimuth journal non-flatness. The yoke will have a pintle bearing at the bottom center and this bearing will constrain the yoke against horizontal translation. This will allow rotation about a vertical axis (azimuth axis). The other side of the pintle bearing will be attached to the center of the telescope pier.

7.5 Telescope Performance

The telescope must meet a number of performance requirements: stiffness, load carrying capacity, allowed displacement tolerances, and motion control. Probably the most difficult requirements are stiffness against gravity and wind loads, and related dynamic performance. The design and performance of the integrated tube and yoke is described by Medwadowski (2002b). A side view of the telescope is given in Figure 7-9. A front view is shown in Figure 7-10, and a plan view is shown in Figure 7-11. An isometric view is shown in Figure 7-12. The masses of the telescope components are given in Table 7-1

7.5.1 Static Deflections Against Gravity

Since the telescope moves in elevation angle, the telescope must adequately hold the optics in proper alignment over the 65° range of zenith angles. It is impractical to hold the optics passively to the required tolerances. Thus, we expect that active control will be needed. The static response of the structure will dictate the range and type of active alignment control that is needed.

The first class of requirement is that the passive support of the optics should be good enough to keep the optics within the active range of adjustment. Thus the structure dictates the range of actuators that support the primary mirror segments, the actuators that control the secondary mirror position, and possibly also the system that controls the tertiary.

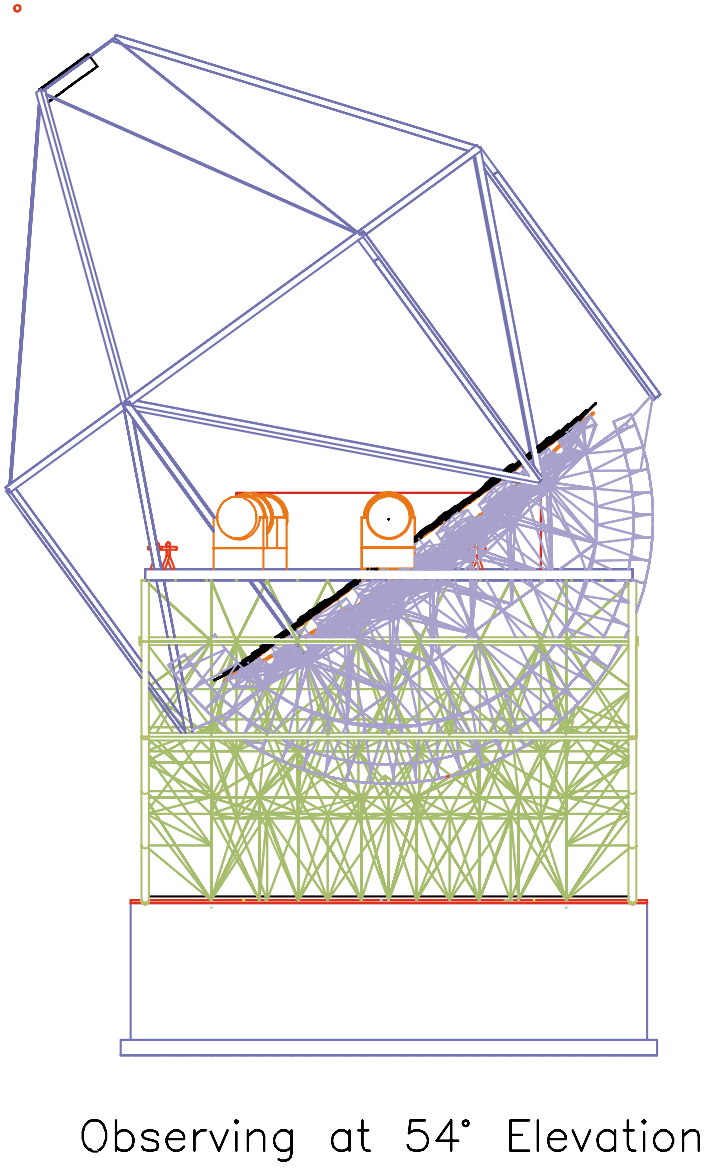


Figure 7-9. The telescope tube and yoke on the pier is shown from the side with the telescope at a zenith angle of 36°.

Axial displacements

As the telescope moves from zenith to 65° the primary will displace along the optical axis (z) by +10.7 mm. Analysis of the structure also shows that the secondary will displace by +15.1 mm along z in going from the zenith to 65°. Thus there is a net increase in the primary-secondary separation in going from the zenith to 65° of 4.4 mm. We expect that this correction can be done open-loop, once a lookup table of corrections has been measured. From Section 4.4.2 we see that 1 mm of axial error introduces 3.09 arcsec of defocus image blur.

In addition to axial displacements of the primary and secondary as rigid bodies, the individual segments and clusters will displace differentially. This has been analyzed in Section 7.2.4, and differential motions of segments are up to 10 mm. This can be reduced by removing the best fitting plane. The minimum actuator range for the segment actuators is set as 2.25 mm for the present design.

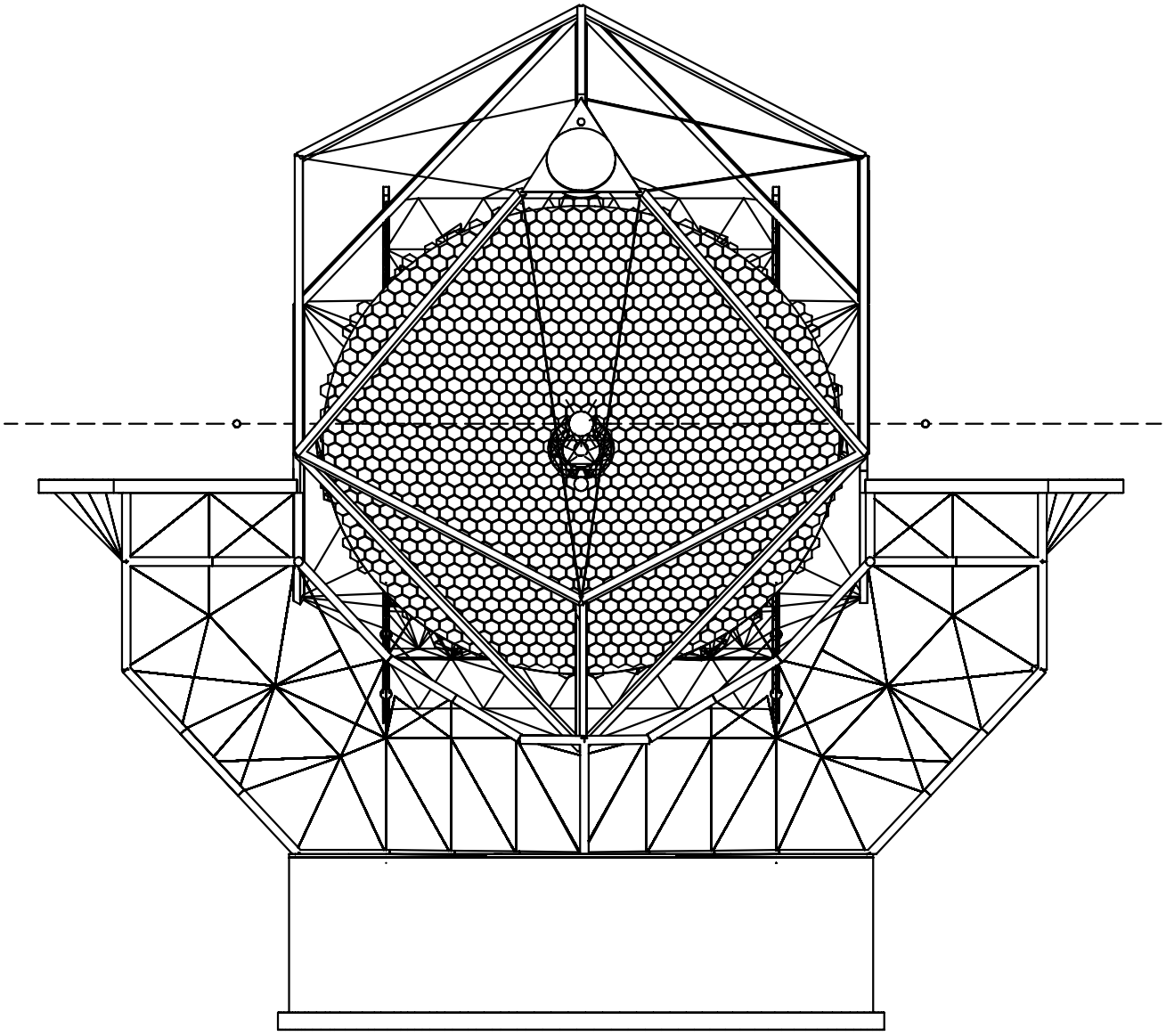


Figure 7-10. The telescope tube, yoke, and pier are shown from the front of the telescope. The telescope is at a zenith of 65°.

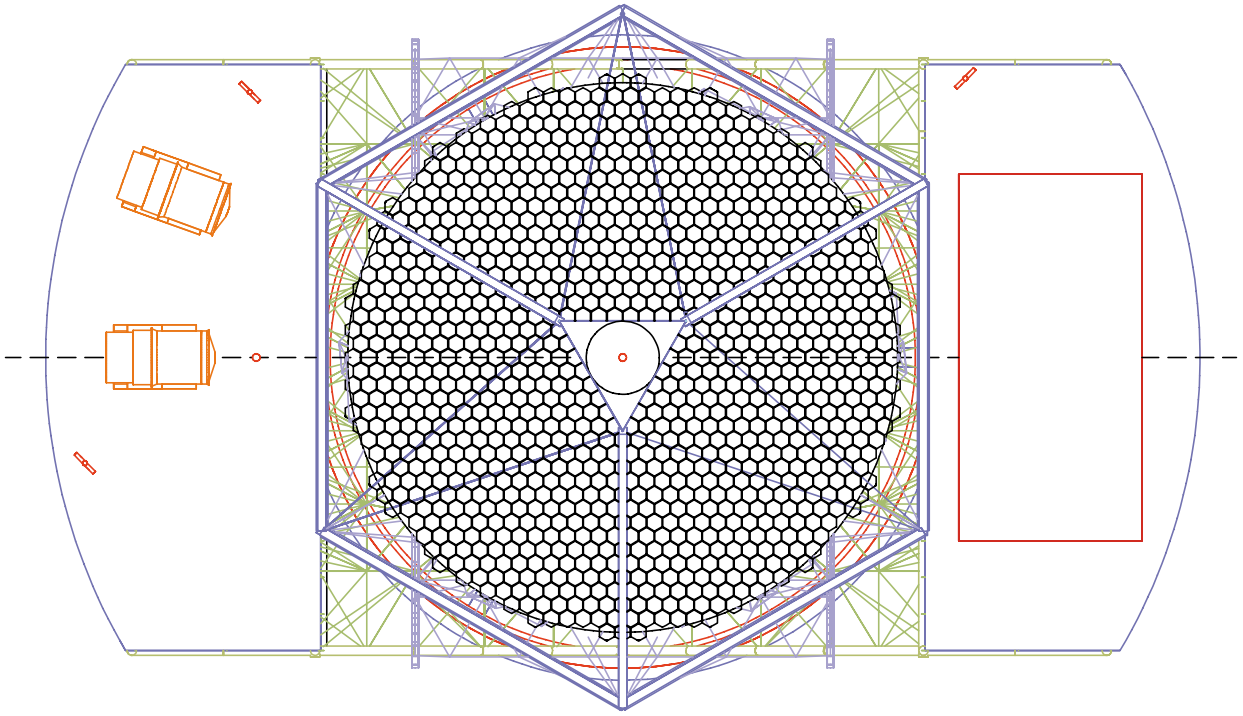


Figure 7-11. This plan view of the telescope shows the upper tube with its blockage of the primary and the Nasmyth platforms with typical instruments placed on them.

Table 7.1 Masses of key components

Component	**Mass of Telescope (tons)**		
Tube assembly	736		
Secondary system		10	
Secondary mirror			3.2
Secondary support			6.8
Upper Tube structure		152	
Tertiary mirror system		10	
Tertiary mirror			2.7
Tertiary mirror support			7.3
Primary Mirror system		154	
Primary Mirror			80
Passive and active support			20
Subcells			9
Clusters			45
Mirror cell and journals		410	
Yoke	440		
Nasmyth scientific instruments	160		
Hydrostatic bearings, drives	20		
Total moving mass	**1356**		

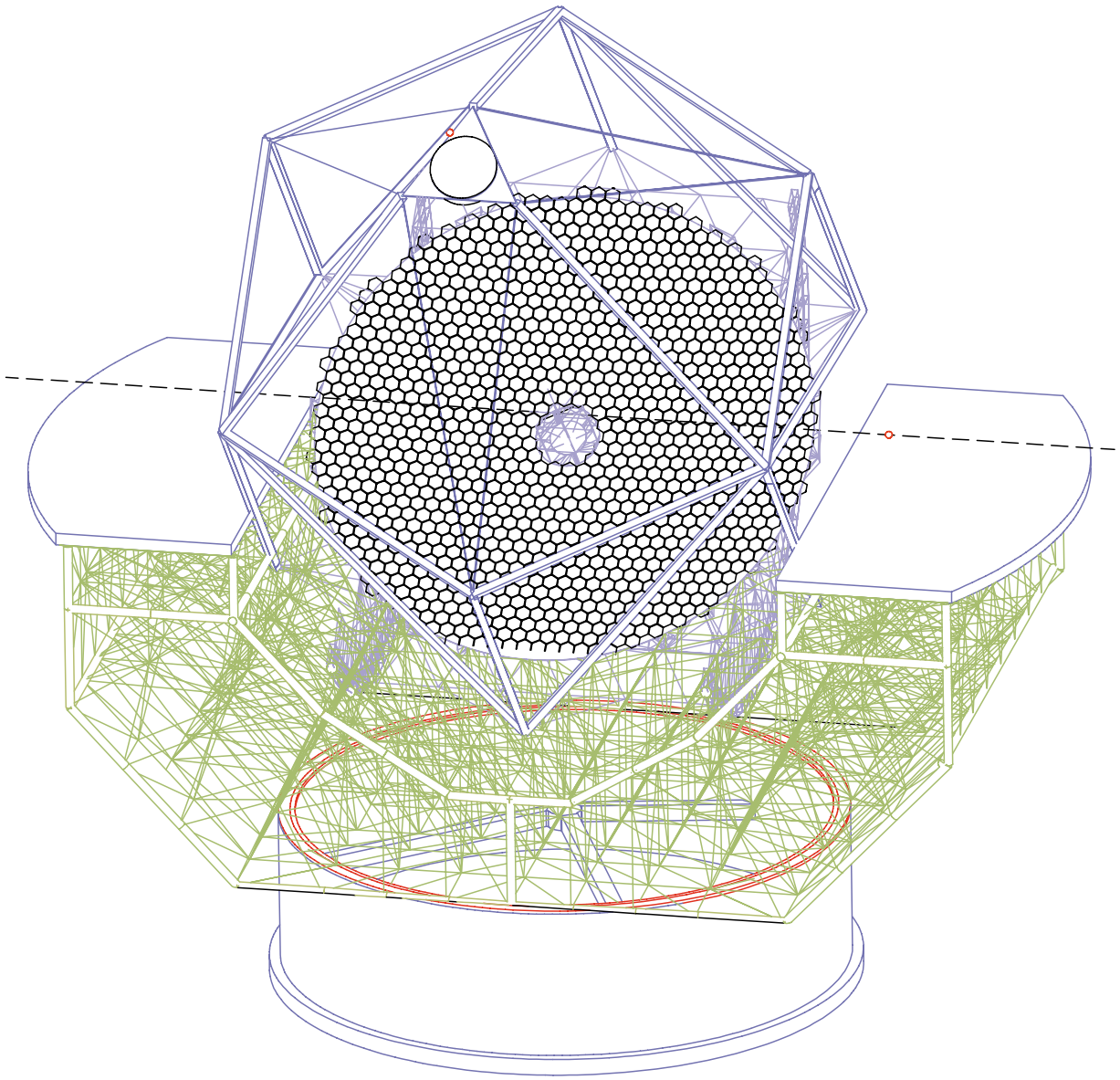


Figure 7-12. The telescope on its pier is shown, with the telescope at 30° zenith angle.

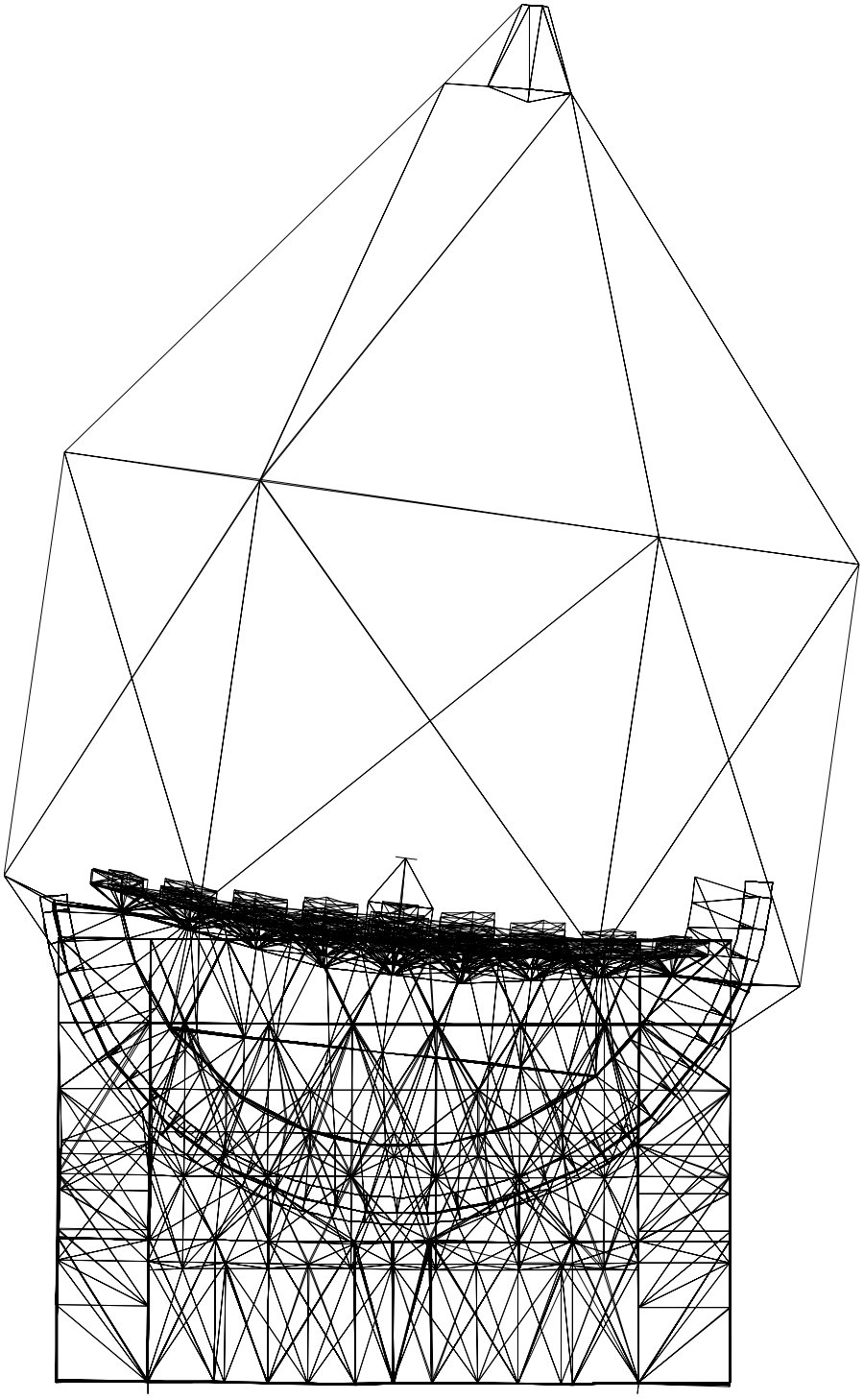


Figure 7-13. The shape of the lowest frequency mode of the telescope is shown. The natural frequency is 1.58 Hz. The amplitude is greatly enlarged to show the shape, which approximates a rotation about the elevation axis.

Lateral displacements
The decenter of the secondary relative to the primary must be controlled to a high level. As mentioned in Section 7.2.4, finite element analysis indicates that the primary mirror will laterally displace by 10.7 mm in going from the zenith to 65° zenith angle. In addition the primary will rotate by -1.09×10^{-3} radians about the x (elevation) axis in going from the zenith to 65°. This leads to a displacement of the primary optical axis at the nominal secondary position of 58.1 mm. However, the gravity driven lateral displacement of the secondary due to structural deformation is 83.1 mm, leading to a net primary-secondary decenter at 65° of 25.0 mm. We saw in Section 4.4.2 that the primary-secondary system must be in lateral alignment to about 0.1 mm. Hence the active control of the secondary position must correct the 25.0 mm to this level of accuracy. We expect that mapping out these changes and subsequently applying them open-loop (with no direct optical measurement) will be adequate, since gravity flexure is quite repeatable. At Keck gravity flexure has been measured with non-repeatability no worse than 1 part in 1000. The autoguider will also in all likelihood have wavefront sensing that can at least measure the misalignment coma and thus ensure it is cancelled closed-loop.

Relative tilts
We have indicated that the primary will rotate (tilt) by -1.09×10^{-3} radians in moving from zenith to 65°. Analysis also shows that the secondary mirror will rotate by -6.1×10^{-4} radians leading to a net secondary tilt of 4.8×10^{-4} radians. In Section 4.4.2 we saw that tilt of the secondary introduces coma at 0.010 arcsec/arcsec. Hence this rotation will introduce 1.02 arcsec of coma. We expect that we will actively correct all secondary position and angle errors.

These rigid body motions will be included in the telescope pointing system, and also in the primary-secondary alignment system.

The relatively large static deflections of the properly aligned primary-secondary system will cause the images to move in the focal plane, but these will automatically be removed by the guiding system. In addition, translations of the primary-secondary optical system relative to the elevation axis can cause a tilt of the focal plane. Assuming a displacement of 20 mm, and assuming the tertiary is tilted to stabilize the image at the instrument, we will tilt the focal plane by the ratio of the displacement divided by the distance to the tertiary, or 0.020/20 = 0.001. At the edge of the 20-arcmin field of view, this will cause a defocus of 0.04 arcsec, an acceptably small value.

A second class of static misalignments is connected to the in-plane segment misalignments. As mentioned in Section 7.2.4, gravity will cause the segments to both translate and rotate in the coordinate system of the primary mirror. The gravity displacements must not exceed the error budget given in Chapter 11. If these exceed the requirements we have several options. We must either introduce active control to these motions: modify the structural design to reduce them, modify the optical design (smaller segments, longer primary focal length) to relax the tolerances, or use warping harnesses to actively change the shape of the segments as a function of elevation angle. As stated in Section 7.2.4, the present design is adequate in this regard.

7.5.2 Static Deflections Against Wind

Wind loads on the enclosure and telescope are a potential problem on windy nights. The details of these wind loads will be site dependent and enclosure design dependent, so quantitative results are not available at this time. Further, the wind loads will have a significant dynamic component; hence static

estimates of the effect of wind loads do not adequately define the problem. Nonetheless, static estimates give one a useful basis for understanding the magnitude of the problem.

The shielding value of the enclosure is extremely important. For Keck Observatory, wind tunnel tests were made of the dome and shutter. The wind speed reduction was measured at a variety of locations within the interior of the dome for a variety of dome elevation and azimuth values. Kiceniuk and Potter (1986) give the detailed reports of this study. The Keck dome geometry is very close to that of the CELT enclosure design (see Chapter 12); the Reynolds numbers are very high in the wind tunnel test ($Re>10^6$) so the measurements should be applicable to CELT.

The measurements are made at a range of distances from the dome center. For CELT the secondary mirror is located at a distance of 35.9 m from the dome center, or 0.798 of the dome outside radius of 45 m. Kiceniuk and Potter made many measurements at a radius of 0.815 of the dome outside radius, so we will use these results herein. They find that at this radius the interior air circulates in a rotational sense, i.e., the air motion is tangential to the dome surface. Further, the absolute worst case they found among all telescope orientations was that the inside air speed was reduced to 30% of the outside free air speed. In all downwind directions, the highest wind speed was 11% of the free air speed. Other measurements were made at a variety of radii. It was found that the residual interior velocity is a rapid function of how deep in the dome one makes the measurement. At the equivalent CELT radius of 40.0 m they find the dome shielding is less effective, with a worst case wind reduction to about 0.46 of the outside free air speed.

There is anecdotal experience at Keck Observatory that is consistent with these wind tunnel results. Observers have complained about telescope shake on very windy nights when pointing upstream into the wind. However, the normal Keck shutter configuration has the shutters fully open. When the shutters were closed to the minimum opening, indications of telescope shake from wind disappeared. From this we can conclude that setting the shutter to its minimum opening (the configuration of the wind tunnel tests) has a significant shielding effect from the wind. More recent tests of the effects of wind have been undertaken for the Gemini enclosure in Chile (Gemini 2000). Similar attenuation factors were found.

Given this data, we can make some rough estimates of wind loads on CELT. We make a very conservative assumption that the wind reduction at the top end of CELT will be 0.40. In reality, observers typically work at orientations that are random with respect to the wind direction; thus one might average the wind reduction factors instead of picking the worst case. We have chosen an attenuation value between the above stated 0.30 and 0.46 to account for the fact that the secondary structure will have a range between 35.9m and 37.9m.

We select the wind statistics of Mauna Kea to work out an example. At Mauna Kea the wind is in excess of 12m/s 10% of the time (Nelson, et al., 1985). Thus we will assume that in these wind conditions the wind speed at the top end of the telescope is 5m/s. Further, we assume that the wind speed drops linearly to zero at the center of the dome. The wind tunnel data indicate that the wind speed drops more rapidly than linear as one goes inwards from the dome edge, hence this is also a conservative assumption.

The static force on the telescope is given by

$$F = \rho v^2 C_D A / 2 \qquad (7\text{-}4)$$

where $\rho = 0.764$ kg/m^3 the density of air at the site (assumed Mauna Kea)
$v = 5$m/s the wind at the top of the telescope
$C_D = 1$ the drag coefficient of the wind on the structural element (typical)
A = cross section of the element.

The drag coefficient is a function of the shape of the element and the Reynolds number (Re) of the airflow past the element. For flat plates $C_D = 2$ is reasonable, but for cylinders it is roughly 1 for Re <250000 and 0.5 for Re > 250000. Thus the shape of the elements can have a significant impact on the net wind force.

For our example, the telescope at the top end will experience static wind pressures of 9.5 N/m^2 (assuming C_D =1).

To estimate the impact on the telescope, it is more relevant to calculate the net torque on the telescope, given the variable wind speed as a function of height along the telescope. We can then calculate the force at the top end that provides this torque, and the secondary mirror motion, and the resulting image motion. This takes advantage of the numerical results of the deflections of our telescope model given a point load at the secondary.

Using the upper tube model described by Medwadowski (2001b), we calculate the effective cross sectional area (weighted by the wind speed squared and the moment arm) to be 21.7m^2. The Medwadowski telescope model (2002b) produces an effective spring constant for lateral displacements at the secondary of 1.04 x 10^7 N/m. Combining the pressure, the effective area and the spring constant yields a static wind deflection at the secondary of 1.99 x 10^{-5} m. Using the decenter-generated image motion arrived at in Chapter 4 (4.13e x 10^3 arcsec/m), we get the final estimated image motion from a static wind of 0.082 arcsec.

This image motion is a rough estimate of what might happen for the 10% external wind, not the average image motion due to wind. Thus, unless dynamics grossly alter this result, we believe this static stiffness will be acceptable. We also believe that with additional design effort the structure can become both stiffer and have a higher natural frequency. Padin and MacMartin (2001) have carried out an independent wind assessment that is consistent with the above discussion.

7.5.3 Dynamics

The dynamic response of the telescope structure is important for several reasons. First, it is a convenient metric for the stiffness of the telescope against various loads, from wind to earthquakes. Second, the lowest natural frequencies will dictate how rapidly the telescope can be moved and controlled. Third, the lowest natural frequencies will dictate how quickly we can move the secondary mirror to remove unwanted image motion. Finally, the dynamic coupling of the wind or other seismic disturbances (vibration) to the telescope depends critically on the natural frequencies, and if sufficiently excited, telescope vibrations will be detrimental to image quality. In all respects, we want a telescope with the highest practical natural frequencies.

The present design of the telescope structure has dynamics that depend sensitively on the elevation angle of the telescope. For a wide range of elevation angles (el > 40°) from the zenith, the behavior is essentially the same as at the zenith. For this range, the lowest natural frequency is about 1.6 Hz. As the zenith angle nears the 65° limit, the performance degrades significantly; and the lowest natural frequency drops to about 1.2 Hz. The lowest modes appear to involve the elevation journals in a significant way, approximating a rocking motion in the y-z plane. The x axis is along the elevation axis. Figure 7-13 shows the distortion in the lowest mode at the zenith.

The lowest mode at the zenith has a natural frequency of 1.58 Hz, and the shape corresponds approximately to rotation about the x axis, with the deformations largely those of the elevation journals. The second mode occurs at 1.85 Hz and appears largely as rotation about the y axis (rocking motion in the x-z plane), with shearing of the journal-mirror support structure. At 65° the mode shapes are similar to the zenith, but the frequencies are 1.16 Hz and 1.59 Hz.

In the next phase of work we will study more carefully the reasons for these low natural frequencies with the hope of improving them. They may be a fundamental consequence of the large elevation journals.

7.6 Pier

The yoke is supported on hydrostatic bearings that move on an azimuth journal. This journal radius is 16.26 m. Because of the large size of the telescope, we see little benefit in elevating the telescope. As a result the azimuth journal will be about 1 m above the ground, a height convenient for maintenance. The total mass of the tube, the yoke, and the scientific instruments will be 1356 tons, and is carried by four azimuth pads to the journal at the top surface of the pier. The pier will be rigidly connected to the earth, relying largely on the mass of the pier. We expect that the pier mass will be about 10 times the mass of the telescope, or 10^4 tons.

The pier shape is shown in Figure 7-10. It is roughly an inverted mushroom with a stiff base disk and a cylinder attached to it. The cylinder depth will be set by the soil conditions, assumed here to be 10 m. A wall thickness of 0.7 m should be adequate. The interior of the pier will be filled with sand or some other dissipative material. The concrete mass will be approximately 3715 tons, and the interior fill will have a mass of roughly 15000 tons.

In Section 7.5.2 we saw that wind loads on the top end of the telescope could be an important issue. In addition to deforming the telescope structure, wind loads will exert a torque on the pier and in turn on the soil under the pier. To estimate the magnitude of this effect we assume we can treat this situation as a disk on a semi-infinite elastic foundation (Richart, et al., 1970).

If the telescope were pointing to the horizon, there will be a rotation about a vertical axis that is given by

$$\theta_s = 3\,\tau_\theta / (16 G r_0^3) \qquad (7\text{-}5)$$

where τ_θ = torque about vertical axis = 7380 Nm
G = soil shear modulus = 4.82×10^7 N/m^2 (for Mauna Kea, Harding-Lawson, 1986)
r_0 = radius of pier = 17 m
yielding θ_s = 0.0012 arcsec

The other extreme is when the telescope is pointing to the zenith and the foundation rotates about a horizontal axis. In this case

$$\theta_s = 3(1 - \nu)\, \tau_\theta / (8Gr_0^3) \qquad (7\text{-}6)$$

where ν = Poisson ratio of soil = 0.3
τ_θ = 12510 Nm
yielding θ_s = 0.0029 arcsec

We conclude that the additional influence of the wind on the telescope rocking the telescope pier is small. Nelson (1983a) and Medwadowski (1984) describe different analytical approaches to this issue.

7.7 Bearings

The telescope motion is most efficiently handled using hydrostatic bearings for both axes. Hydrostatic bearings support very large loads, have extremely small drag and stiction, and are extremely stiff. Their major concerns are the necessity of handling liquid oil and avoiding consequent vibrations introduced by the oil pumps. A conceptual study of CELT bearing options and the characteristics of a hydrostatic system in particular was carried out by Vertex RSI (2001).

The vertical load on the elevation bearings is 736 tons, and with an opening angle of 50°, each of four bearing pads must carry a load of 1.99 x10^6 N. With six elevation bearings the load distribution will be different, but we will not explore the details here. Assuming that the hydrostatic bearing oil is supplied at a pressure of 1.0 x10^7 N/m^2 (1450 psi) each of the elevation pads needs an effective area of 0.20 m^2. The four azimuth pads carry a total load of 1356 tons, thus each pad must carry a load of 3.33 x10^6 N. Assuming the same oil pressure as the elevation pads, each azimuth pad needs an effective area of 0.33 m^2.

As the telescope moves, oil must be supplied at a sufficient rate that the pads do not overrun the oil during slewing. To meet our slew requirements (>1.2°/s az, 0.22°/s el), the pads must move relative to the journal at up to 0.36 m/s (0.064 m/s el). This implies the flow rate to the set of pads (assuming an oil thickness of 50 μm) must be at least 1.42 x 10^{-4} m^3/s (2.14 x 10^{-5} m^3/s el). The total oil flow is thus 1.63 x 10^{-4} m^3/s (9.8L/minute).

The hydrostatic oil heats up as it exits the pad ($\Delta T = p/c\rho$). To avoid adding heat to the dome volume, the oil should be precooled by about 7.0°C. This will allow the oil to exit at the ambient temperature. The thermal power is about 1.6 kW.

One of the virtues of hydrostatic bearings is their extreme stiffness. With a film thickness of 50 μm, the stiffness is 6 x 10^{10} N/m, orders of magnitude stiffer than the telescope structure. Pumps and cooling systems with these capacities are routinely available, so we foresee no unusual problems with the use of hydrostatic bearings on CELT.

As mentioned earlier, the lateral restraint of the yoke will be a pintle bearing at the center of the yoke. This bearing will be a rolling element bearing and by design will only carry horizontal loads. In elevation, motion along the elevation axis will be constrained by additional hydrostatic bearings, a technique used successfully on Keck.

7.8 Drives and Encoders

The drive system for the telescope must meet the slew requirements and also move the telescope smoothly enough for scientific observations. Vertex RSI (2001) studied several drive systems including wheel and track, direct drive (used on the VLT), and gear driven drives. They recommend that a gear driven system can economically meet our requirements. To meet our slewing requirements (acceleration requirements) they recommend 4 x 37 hp motors in azimuth, and 4 x 8 hp motors in elevation.

Geared systems may introduce ripple into the response, which is undesirable. However, with pairs of motors opposing each other, hysteresis and backlash can be prevented, and with modern controls and suitable encoding, smooth performance can be achieved. Helical gears are suggested as excellent candidate gear systems because of their relatively uniform contact area. Helical gears are being used on Southern Astrophysical Research Telescope (SOAR, 2001), with excellent results.

In azimuth the gear is likely to be part of the pier/journal system while two pairs of opposed motors will be attached to the yoke. In elevation the gears will be attached to the two journals, and the two pairs of motors will be attached to the yoke. Dynamic analysis by Vertex RSI indicates that all the dynamic motion requirements can be met with such a system.

Various encoder systems can be envisioned. Again, Vertex RSI explored several encoder options, and there were several possible solutions including Heidenhain tapes, Farrand-Inducosyn, and BEI absolute encoders. Different systems have different strengths and weaknesses, but adequate encoders do exist for CELT. Other less conventional systems such as inertial reference units may also have a useful role in CELT. In azimuth there is a complete circle and also a pintle bearing, so the encoding issues are straightforward. In elevation there is only an arc of a circle; hence the encoding is potentially more difficult. Vertex RSI has suggested that the use of Heidenhain tapes (common on other telescopes) would be sufficient for our needs. These tapes give ample resolution, providing up to 0.0004 arcsec resolution (but deformations of the journals might limit their accuracy).

7.9 Thermal Responses

The telescope system must work successfully in varying thermal environments. Such variations include a wide range of static temperatures, diurnal variations, and local temperature variations across the structure. Variations in the average temperature from –2°C to +6°C (for Mauna Kea, this covers 90% of the time) should not cause any difficulties.

As mentioned at the beginning of this chapter, temperature change, temperature gradients, and more complex temperature variations will cause deformation of the structure and motions of the supported optics. These are critical issues to be explored more fully in the next stage of design. The rough estimates made currently suggest that with the planned active control of the optics, these effects will not be a problem.

The telescope structure also has significant heat capacity; hence its temperature will not perfectly follow the ambient air temperature inside the dome. These temperature differences will produce some temperature variations in the surrounding air, and thus degrade the seeing. At 1.5×10^6 kg, the telescope structure has a heat capacity of 7.05×10^8 J/°C and must dissipate 195 kW to cool by 1°C in one hour. The entire dome air mass is only ~ 3×10^5 kg with a heat capacity of 3×10^8 J/°C. This indicates a central issue to be addressed in the next design stage: how one causes the telescope structure to follow

the changing ambient air temperature or isolate the thermal mass of the structure, so that its skin temperature stays within about 0.5°C of the ambient air temperature.

7.10 Instrument Changing System

We plan to locate all scientific instruments on the Nasmyth platforms. With an articulated tertiary, the science light will be directed to the chosen instrument by moving the tertiary. Thus, there will be no need to routinely transport the instruments. We expect that the instruments will be mounted/assembled on the Nasmyth platforms and will be permanently stationed there. Servicing will be done in place, or key components will be removed and transported to laboratory space in the support building.

An elevator will be available to transport moderate-sized objects in support of instruments. The location of the elevator will be determined in the next phase of design. It may be stationary, as at Keck, and the telescope must be positioned suitably to allow transfer, or it might be a part of the telescope itself.

More serious instrument installation and handling will be by cranes. We expect that the dome will have a 5-ton crane, and expect that a mobile crane with over 20-ton lift will also be available at the observatory. This should be sufficient for instrument assembly and major servicing.

7.11 Segment Handling

A critical part of CELT is its 1080-segment primary mirror. We will need to install and align these segments, regularly clean them in place, and periodically remove them to be re-coated. With this many segments these processes must be carefully understood and optimized to avoid excessive effort, loss of night time, or the introduction of any significant risks to the segments themselves.

As mentioned earlier, we plan to place segments in clusters of 19, and the segment relative alignment will be set within the cluster. Thus a facility in the support building will be needed to assemble, align, and disassemble individual clusters.

Segment cleaning in place will likely be done with CO_2, which has worked well at Keck. At Keck this is done manually from a crane. At CELT, a more automated system is desired and will be developed in the next phase of work.

Cluster installation and removal is still a major issue. At Keck, segments are handled individually with a custom crane, but it has proven to be time consuming and also causes the loss of some night time. For CELT we need to develop a more convenient system, and this will be a major task in the next design phase. We expect that the wavefront sensing system for CELT will provide rapid realignment of segments by making more measurements in parallel, so that nighttime losses will be minimal.

7.12 IR Design Considerations

In the thermal IR wavelength region (> 2 μm), the emission of the telescope itself can be a significant source of background light to the scientific instruments. Proper baffling of the instrument will greatly reduce this, and only the upper tube structure is likely to be of key importance, since parts of it are directly in the optical path. When appropriate for the instrument, these parts can be blocked by a cold pupil stop in the instrument. By keeping the blockage to a minimum, we expect the structure will not be a major background source in the IR. It appears that the segment gaps and edges are likely to be a

difficult-to-remove source of background, a phenomenon fundamental to segmented mirror telescopes. As mentioned in Chapter 4, the segment gaps and edges constitute about 0.9% of the primary mirror area.

7.13 Control of Secondary and Tertiary

In Chapter 6 we discussed the secondary and tertiary controls. We plan to actively control the rigid body motion of the secondary. In addition, the secondary will have an active support system that can change the secondary shape. We expect this will be used to set and stabilize the secondary surface shape. The tertiary will have active control over two rigid body motions, rotations about the z axis, and rotation about the local y axis. None of these controlled motions are expected to have significant dynamic impact on the telescope structure, and although the attachments have not been worked out, we expect this will be a straightforward interface.

We want to actively control the secondary mirror to reduce undesired image motion (windshake) and to chop the secondary to improve IR performance. These will be engineering challenges: rapidly and accurately tilting a 10-ton object, and minimizing the dynamic interactions with the rest of the telescope. In the next phase of work we will address these challenges. Engineers responsible for large chopping secondary mirrors on Keck and Gemini have expressed optimism that this is practical at the 1 Hz level (Lorell 2001).

7.14 Field Rotation and Other Effects of Alt-Az Mount

An altitude-azimuth telescope has structural advantages, but it adds an astronomical complication. The scientific field of view rotates relative to the telescope as the telescope motion removes the star motion caused by the earth rotation. This field rotation is generally harmless when a single star is being studied, but when an extended object or region is being studied, this rotation must be removed. In addition, the telescope motions in azimuth become infinitely rapid as the observing region approaches the zenith. Thus there is an effective blind spot near the zenith. More generally, the motions of the telescope about the elevation and azimuth axes are variable, even though the Earth's rotation rate is constant.

Nelson (1981) describes the quantitative aspects of telescope and field rotation. The additional motion of the tertiary needed to keep a star image on a given instrument at any location on the Nasmyth platform is described by Kuhlen (2001). There are no singularities or other unusual motion requirements of the tertiary.

REFERENCES

Gemini South 8-m Optical Telescope, Final Report. Modal Analysis and Controls Laboratory Report No. -05-08570-001. CD-ROM produced by David R. Smith: "Wind Tests Data from Gemini South." May 13-14, 2000. Mechanical Engineering Department, University of Massachusetts: Lowell, Massachusetts.

Harding Lawson Associates. 1986. "Soil Investigation Keck Observatory 10-m Telescope, Mauna Kea, Hawaii." Keck Observatory Report No. 152.

Kiceniuk, Taras and Kent Potter. 1986. “Internal Air Flow Patterns for the Keck 10 Meter Telescope Observatory Dome.” Keck Observatory Report No. 166.

Kuhlen, Michael. 2001. “The Tertiary Mirror Equations of Motion of an “off-altitude axis” Nasmyth Focus.” CELT Report No. 22.

Lorell, Ken, 2001, private communication.

Medwadowski, Stefan J. 1984. “UC telescope pier rotations due to wind action on the observatory dome.” Keck Observatory Report No. 53.

Medwadowski, Stefan J. 2001a. “The structure supporting a cluster of segments of the primary mirror of CELT.” CELT Report No. 25.

Medwadowski, Stefan J. 2001b. “The structure of the tube of the CELT telescope.” CELT Report No. 26.

Medwadowski, Stefan J. 2001c. “CELT tube structure with upper tube in the form of a tripod-on-hexagon.” CELT Report No. 27.

Medwadowski, Stefan J. 2001d. “CELT Tube Structure, Final Conceptual Design.” CELT Report No. 28.

Medwadowski, Stefan J. 2002a. “CELT Yoke Structure, Final Conceptual Design.” CELT Report No. 29.

Medwadowski, Stefan J. 2002b. “CELT Structure (Tube on Yoke), Final Conceptual Design.” CELT Report No. 30.

Meinel, Aden and Marjorie Meinel. 2000. “Configuration Options CELT.” CELT Report No. 31.

Nelson, Jerry. 1981. “Geometrical Relations for an Altitude Azimuth Telescope.” Keck Report No. 49.

Nelson, Jerry. 1983a. “Telescope pier rotation from wind loads on building and dome.” Keck Technical Note No. 58.

Nelson, Jerry. 1983b. “The effects of wind loads on the ten meter telescope.” Keck Report No. 47.

Nelson, Mast, and Faber. 1985. “The Design of the Keck Observatory and Telescope.” Keck Observatory Report No. 90.

Padin, Steve and Doug MacMartin. 2001. “Aberrations due to wind-induced secondary decenter in CELT.” CELT Technical Note No. 24

Richart, F. E., J.R. Hall, Jr., R. D.Woods. 1970. *Vibrations of Soils and Foundations*, Prentice-Hall.

SOAR, http://www.soartelescope.org/, 2001.

Vertex RSI. October 2001. “30-m CELT Telescope Drives, Bearings and Encoder Study.” CELT Report 21.

von Hoerner, Sebastian. 1967. *Astronomical Journal* **72**, 35.

Chapter 8. Optics Alignment

8.1 Introduction

Aligning the CELT mirror segments and co-aligning the primary and secondary mirrors is similar to aligning the Keck optics. The major differences are that the number of degrees of freedom (which scales with number of segments) is a factor of thirty higher for CELT, and the tolerances are a factor of several tighter. Nonetheless, we are confident that there are no fundamental technological barriers to be overcome in extending the highly successful Keck alignment procedures -- based largely on the Phasing Camera System (PCS) -- to CELT. Nominally, the higher number of degrees of freedom would require detectors that are five to six times larger; however, the Keck PCS CCDs (1024 x 1024) were not used efficiently in this sense. Since 2048 x 2048 CCDs are readily available, we expect it will be straightforward to implement a similar method.

The increase in the number of degrees of freedom will necessitate somewhat more robust algorithms. Statistically, one could expect to encounter a 3 σ measurement error about once in every three Keck alignment procedures; for CELT, about *ten* such errors will occur during every procedure. We will devote substantial effort in the coming design phase to investigating these issues, largely through numerical simulations, supplemented by Keck and other laboratory data.

The principal deficiency of the Keck alignment and wavefront control system is, arguably, the relatively poor control of telescope focus (secondary despace and also, possibly, focus mode of the primary mirror). For CELT this will be remedied by incorporating focus measurements (and perhaps other low spatial frequency primary mirror modes) into enhanced guiders or on-instrument wavefront sensors, so that these modes will be under closed-loop control.

8.2 Degrees of Freedom

We begin the discussion of the alignment of the CELT optics by reviewing the techniques used to align the optics of the Keck telescopes. A long and substantial investment in the hardware and software used to align the Keck telescope segmented primary, secondary, and tertiary mirrors has provided us with a wealth of information that can be readily applied to CELT and will save substantial costs and time.

8.2.1 Keck Degrees of Freedom

Each Keck telescope primary has 36 segments with a total of 105 out-of-plane segment rigid body degrees of freedom, plus 3 overall primary rigid body degrees of freedom. The other 108 rigid body degrees of freedom involve in-surface motions. With their looser tolerances, they are set by the initial surveying during installation. Although each segment has warping harnesses that can adjust 30 degrees of freedom describing the surface shape, in practice only the three quadratics (two segment astigmatism terms plus segment focus) are significantly controlled. Three actuators control the three out-of-plane degrees of freedom of the secondary mirror. Only one degree of freedom of the tertiary, rotation about the optic axis, is controlled and that only in discrete steps defined by detents. Thus there are about 220 degrees of freedom regularly adjusted at each Keck telescope.

8.2.2 CELT Degrees of Freedom

CELT will have 1080 segments with a total of 3237 out-of-plane segment rigid body degrees of freedom plus three overall primary rigid body degrees of freedom. Possibly three (non-rigid body) degrees of freedom of each segment surface will also need to be controlled. We will control five of the six rigid

body degrees of freedom of the secondary. (An active secondary, if adopted, would involve perhaps another 30 to 50 degrees of freedom associated with the secondary mirror shape.) The CELT tertiary will be controlled in two rotational degrees of freedom in order to use off-axis Nasmyth instruments. Thus there will be about 6500 degrees of freedom to be actively controlled at CELT. This factor of thirty increase in controlled degrees of freedom will be a challenge (particularly with regard to the complexity and required robustness of the associated algorithms), but one that can be met based on the information gained through the long and rich Keck experience.

8.3 Optics Alignment Experience at Keck and Implications for CELT

A substantial amount of the alignment experience at Keck is described by Gary Chanan and his group (Chanan 1994a,1994b,1998,1999, 2000). We give here an overview of that work.

8.3.1 Alignment Tools

The Keck telescopes make use of four alignment tools to measure the above degrees of freedom.

Phasing Camera System
The Phasing Camera System (PCS) is a Shack-Hartmann sensor that operates in four separate modes:

Passive Tilt mode measures segment tip/tilt angles by utilizing one Shack-Hartmann subaperture per segment.

Fine Screen mode measures segment tip/tilt angles as well as second order segment aberrations. The latter are used to infer the position of the secondary. There are 13 subapertures per segment in Fine Screen, although the central subaperture is usually ignored because of systematic effects associated with the central dimples (roughly 15 cm in diameter and 400 nm deep) on the Keck segments.

Ultra Fine Screen mode puts 217 subapertures over a single segment in order to measure the segment shape for warping harness adjustments. UFS can only measure one segment at a time. In this mode, unlike Fine Screen, the segment surface is densely sampled. A new mode, still under development, puts 127 subapertures across each of seven segments; all seven can be measured at once.

Phasing mode puts subapertures across each of 78 intersegment edges (6 additional edges are blocked by the secondary mirror and tertiary tower). Phase (segment piston) measurements are made by a physical optics generalization of the usual (geometrical optics) Shack-Hartmann test.

PCS is principally used in two situations: 1) after a segment exchange, when it requires between one half and one full night, depending on how many segments were replaced (with the time requirement driven by the need to use the UFS or segment surface measuring mode), and 2) roughly once a month to correct for sensor drift in the active control system, when it requires 1-2 hours to make all the necessary measurements. Our plans for CELT are to reduce the time required for the first situation by making the segment surface measurements simultaneous with array wavefront measurements, and by automating pupil registration, which is currently done manually at Keck. The time that is required for the second situation would be acceptable for CELT.

We expect that CELT will require all four of the PCS functions, although it might be possible (and it is certainly desirable) to combine the Fine Screen and UFS modes to save costs and complexity. We believe it would be straightforward to build a Keck-style PCS instrument for CELT, and basing the associated software on that developed at Keck can save major development costs. CELT may also utilize some of the following supplemental Keck alignment tools.

MAlign

MAlign ("em-a-line") is a utility that runs on most, but not all, of the Keck instruments on a simple dedicated optical imager known as the Star Stacking Camera. MAlign is an active test in which the positions of the segments are measured by distorting the primary in focus mode (see Section 5.4.1). The stellar subimage formed by each segment is displaced radially from its nominal position. Once well separated, all 36 image centroids can be measured. Since the amplitude of focus mode is well controlled, the expected centroids are known. MAlign cannot distinguish global aberrations caused by the tip, tilt, or despace of the secondary from similar global aberrations of the primary mirror; instead it attributes *all* global focus and coma to a misalignment of the secondary. In terms of the PCS functions, MAlign does not measure segment phase or shape, but does measure segment tip/tilt as well as secondary tip, tilt, and piston where, as noted above, the secondary is positioned to best cancel out global focus and coma in the primary. In adjusting segment tip/tilt, MAlign also constrains the segment pistons to minimize the change in the intersegment edge steps.

MAlign is run at the beginning of every night and possibly a few times during the night, depending on the image quality. The principal corrections are due, or at least attributed, to misalignment of the secondary. A typical MAlign runs takes about three to ten minutes. Ideally, such a procedure will not be required for CELT, since the plan is to control the associated degrees of freedom closed-loop; however, such a utility will still likely be useful for commissioning and trouble-shooting.

Phase Discontinuity Sensing

Phase Discontinuity Sensing (PDS) is a utility that currently runs only on the Near Infrared Camera (NIRC) on Keck 1, although future implementations are planned on other infrared instruments. It deals only with segment phase, and has a much smaller capture range ($\pm$ 400 nm) and somewhat larger errors than does the corresponding mode of PCS (for which the capture range is $\pm$ 30 μm). PDS was developed because the infrared instruments on Keck 1 (which are the only seeing-limited instruments sensitive to segment piston errors) use a different secondary mirror than does PCS. Prior to the development of PDS it was difficult to check or adjust the segment phases once an infrared observing run had started.

PDS is run at the beginning of a NIRC run. Typically only several weeks have elapsed since the last PCS phasing run. A PDS run takes about 45 minutes and must be preceded by an MAlign run. The total time required is about one hour.

Guiders

Guide cameras ("guiders") associated with each instrument monitor the global tip/tilt of the optical system with respect to the target. Guiders run virtually continuously during Keck exposures. Significantly enhanced versions of these guiders, or on-instrument wavefront sensors, will be designed for CELT (see Section 8.3.2).

8.3.2 System Performance

All four of the above alignment-monitoring tools perform well at Keck. The principal shortcoming of the overall system is that none of these tools can perform continuous monitoring of low spatial frequency aberrations. In particular, telescope focus and/or its segmented primary mirror analog, focus mode, and possibly the next few higher-order global aberrations should be monitored continuously. Measurement of these aberrations would both improve image quality and save telescope time, especially with regard to global focus errors. In retrospect, the best way to accomplish this at Keck would have been to upgrade the instrument guiders to be low-order wavefront sensors that could measure not just global tip and tilt, but also at least one (focus) and possibly a half dozen other higher-order aberrations as well.

In Section 8.5 below we discuss the relation between focus and focus mode for CELT, and in Section 8.6 we discuss the low-order wavefront sensors that are used for active optical control of current 8-meter class telescopes. These serve as models for similar enhanced guiders, or telescope control wavefront sensors, for CELT.

8.4 Optical Effects of Motions in Various Degrees of Freedom

As described in Section 8.2, the overall CELT telescope system has a very large number of degrees of freedom: more than 3000 for out-of-plane rigid body motions of the segments, and many thousands more for segment shapes. In addition, five or six degrees of freedom are associated with rigid body motion of the secondary, perhaps another 30 to 50 associated with the secondary mirror shape, and three for the out-of-plane motion of the tertiary. Exercise of virtually any one of these degrees of freedom affects the final image produced by the telescope, and exercise of any of those degrees of freedom not associated with the primary mirror affects the exit pupil.

During the preliminary design phase we will specify in a systematic way the effects that the above degrees of freedom have on the following critical optical parameters: image aberations, image position, image scale, pupil position, pupil scale, focal plane tilt, and distortion.

In the following section we present a useful theorem which can be used to calculate in a straightforward way the image position associated with an arbitrary optical aberration. We also discuss, as a simple example, the image and pupil position shift induced by rotation of the secondary mirror. This is only one example of the systematic analyses that will be carried out in the preliminary design phase.

8.4.1 A Theorem on Image Position

Suppose an optical aberration is completely specified by its expansion in Zernike polynomials. In this circumstance, the position (centroid) of the image can be quickly calculated by means of the following theorem, which is sufficient to state and prove for a single Zernike aberration. At the current level of design, this theorem finds a trivial application in the relation between image position and secondary mirror rotation (see Section 8.4.2). In subsequent design efforts, it can be used to determine, for example, the image shift associated with various global distortions of the primary; and in practice, it can be used to stabilize the image position as the active control system is adjusted.

Let the Zernike aberration in question be Z_{nm}, with coefficient C_{nm}. Here m and n are the azimuthal and radial Zernike indices respectively. We use the definition of Zernike polynomials given by Born and Wolf (1980).

Theorem: A Zernike aberration does not displace the centroid of the image unless m = $\pm$1 (cos θ and sin θ) (and hence n is odd). If m = +1, then the x-centroid displacement will be f C_{mn} /R, and the y-displacement will be zero. If m = -1 , then the y-centroid displacement will be f C_{mn} /R, and the x-displacement will be zero. Here f is the system focal length and R is the radius of the mirror.

Proof: The x and y coordinates of an image centroid are proportional to the x and y components of the wavefront (and hence Zernike) gradient. The Zernike gradient can be expressed as a finite Zernike series over terms of lower order:

$$\partial Z_j / \partial x = g^x_{jj'} Z_{j'} \qquad \text{and} \qquad \partial Z_j / \partial y = g^y_{jj'} Z_{j'} \tag{8-1}$$

where the g^x and g^y are matrices whose elements are constant with dimensions of reciprocal meters and a summation over repeated indices is implied. The index j orders the Zernike polynomials: j = 1, 2, 3...

Since the Z_j are orthogonal, if we integrate Eq. (8-1) over the surface, the only non-vanishing contributions come from terms with j′ = 1 and hence n′ = m′ = 0. One can show (see Noll 1976, but note differences in Noll's Zernike numbering and normalization) that we have:

$$g^x_{jj'} = 1 \ \text{(m = 1 and n odd)} \quad \text{and} \quad g^y_{jj'} = 1 \ \text{(m = -1 and n odd)} \tag{8-2}$$

The matrix elements vanish for all other m and n. The theorem follows directly.

Note that the theorem may be checked trivially for tip and tilt, and, with slightly more effort, for coma.

8.4.2 Image Position Shift as a Function of Secondary Rotation Angle

Rotation of the secondary mirror by an angle ε_x (ε_y) about the x-axis (y-axis) tangent to the secondary vertex will shift the image by angles θ_y (θ_x) in radians on the sky where

$$\theta_y = 2\,\varepsilon_x\,(d + e) / f \qquad \text{and} \qquad \theta_x = -\,2\,\varepsilon_y\,(d + e) / f \tag{8-3}$$

and where d = primary-secondary distance, e = back focal distance, and f = final focal length. The nominal CELT values are d = 39.41 meters, e = 16.50 meters, and f = 450.00 meters, so that the ratio θ / ε = 0.248. This translates into a fairly tight constraint on control of the secondary mirror rotation.

8.4.3 Pupil Position Shift as a Function of Secondary Rotation Angle

We note that mid-infrared instruments will almost certainly need to be equipped with pupil masks to minimize the effects of emissivity of warm structures such as the secondary supports or the dome walls and floor behind the irregularly shaped primary mirror. Thus small misalignments of the exit pupil can have serious consequences for the backgrounds of these instruments.

Rotation of the secondary mirror by an angle ε_x (ε_y) about the x-axis (y-axis) tangent to the secondary vertex will shift the pupil by a linear distance

$$dy = 2\,d\,\varepsilon_x \qquad \text{and} \qquad dx = -2\,d\,\varepsilon_y \tag{8-4}$$

where dy and dx represent the pupil motion referred to the primary mirror. Thus to control the pupil to within 0.1% of its diameter requires control of the secondary mirror tip/tilt to within 57 arcsec. This is a weak constraint compared to the image shift constraint of the previous section.

8.5 Focus and Focus Mode

8.5.1 Introduction

As discussed in Section 8.3, the experience with Keck makes it clear that CELT or its instruments should be equipped with enhanced guiders or on-instrument wavefront sensors, capable of detecting and correcting not just global tip and tilt, but at least one and probably several low spatial frequency modes as well. These should have a moderate bandwidth, 0.01 Hz or higher, where the lower limit is set by the need to integrate over atmospheric turbulence. However, an important issue remains. Low order modes of the segmented primary mirror are almost, but not exactly, degenerate with corresponding global continuous classical aberrations, such as defocus or coma, which can arise from despace or tip/tilt of the secondary mirror. Because of this near degeneracy it may be difficult to distinguish the primary modes from the corresponding continuous aberrations. For example, a Shack-Hartmann sensor cannot distinguish focus from focus mode with only one subaperture per segment. Instead a total of several thousand subapertures would be necessary for this purpose for CELT. On the other hand, this same degeneracy suggests that it may not be necessary to resolve these subtle differences, at least not with high bandwidth. That is, we might use the piston degree of freedom of the secondary mirror to make quasi-real time corrections of both focus (exactly) and focus mode (approximately). At periodic and relatively infrequent intervals, we would make more detailed measurements to resolve the degeneracy and re-zero both aberrations. This is the approach that is currently used at Keck, where the nightly measurements with the MAlign utility do not attempt to distinguish focus and focus mode; the appropriate corrections are made once a month using PCS.

In the following section we present some of the mathematics necessary to make these considerations quantitative; we do not attempt to resolve the question of whether one should detect focus or focus mode at moderate bandwidth. (Although here we restrict ourselves to the focus/focus mode problem, a similar issue arises in the context of other low-order aberrations and modes, e.g., global coma and primary mirror "coma mode," but we defer these latter calculations to the preliminary design phase study.) First, however, we need to clarify some terminology. **Focus mode (FM)** is a global distortion of the primary mirror in which the actuator length changes lie along the surface of a sphere. Equivalently, it corresponds to a constant dihedral angle between all adjacent segments. If singular value decomposition of the control matrix of the active control system is used to define and order all possible modes of the primary mirror (ordering by increasing spatial frequency or decreasing noise multiplier), then Mode One is very close but not identical to focus mode. By **uncompensated focus mode**, we mean that aberration of the overall telescope system consisting only of focus mode in the primary mirror. This is to be distinguished from **compensated** or **differential focus mode**, in which the overall telescope aberration is minimized by canceling focus mode to the maximum extent possible by pistoning the secondary mirror.

8.5.2 Useful Formulae and Conversion Factors

Depending on the context and application, there are at least eight useful ways to characterize the amount of focus mode. Six of these describe uncompensated focus mode:

FM_a = focus mode in microns of rms actuator length change
FM_s = focus mode in microns of sensor offset
FM_z = focus mode in microns of segment Zernike radial tip/tilt for a segment whose center is located at D/2 from the center of the primary
FM_i = focus mode in arcseconds of segment radial image motion on the sky for a segment whose center is located at D/2 from the center of the primary
FM_f = focus mode in microns of primary focal length change
FM_2 = focus mode in microns of equivalent secondary piston

The remaining two relate to compensated or differential focus mode:

FM'_c = focus mode in microns of segment Zernike focus coefficient C_{20}
FM'_r = focus mode in arcseconds of maximum angular deviation of any ray

All of these measures can be expressed in terms of FM_a. If we approximate the overall shape of the primary as a circle, then we can obtain explicit expressions for all of these measures except one. This approximation causes errors that are typically 2-3%. The one exception is the relation between FM_s and FM_a, which is best evaluated numerically and depends on the details of the actuator and sensor geometry. For CELT we assume that the actuator geometry is similar to that of Keck (but scaled appropriately). For the CELT sensors we assume that the sensor locations along an intersegment edge are similar to those of Keck; the geometry is further defined by taking $\alpha = 0.10$, an intermediate value for the CELT sensor design geometry (see Section 5.4.2).

Define the following notation

F_1 = primary mirror F-ratio (1.5)
D = primary mirror diameter (30 m)
f = telescope final focal length (450 m)
a = hexagon side length (0.5 m)
m = magnification of secondary (10)
$\gamma = (1 + m^2)\, a^2 / (2f^2)$ where we have given the reference optic design values in parentheses.

Derivations yield

$$FM_z = 8\sqrt{3}\, a\, FM_a / D \qquad (8\text{-}5)$$

$$FM_i = 8 \times 206265 \times 10^{-6}\ \sqrt{3}\ FM_a / D \qquad (8\text{-}6)$$

$$FM_f = 32\ \sqrt{3}\ F_1^2\, FM_a \qquad (8\text{-}7)$$

$$FM_2 = 16\ \sqrt{3}\, (a/D)^2\, FM_a \qquad (8\text{-}8)$$

$$FM'_c = 4\ \sqrt{3}\, (a/D)^2\, FM_a \qquad (8\text{-}9)$$

$$FM'_r = 16 \times 206265 \times 10^{-6}\ \sqrt{3}\, a\, FM_a / D^2 \qquad (8\text{-}10)$$

Using the optical design parameter values yields:

Table 8-1. Focus mode measures corresponding to FM_a = 1.0 micron

	Keck	CELT
σ	= 1.340 x 10^{-3}	6.23 x 10^{-5}
FM_a	= 1.0 micron	1.0 micron
FM_f	= 170 microns	125 microns
FM_2	= 167 microns	123 microns
FM_z	= 1.25 microns	0.231 microns
FM_i	= 0.286 arcsec	0.095 arcsec
FM_c	= 0.0561 microns	0.00192 microns
FM'_r	= 0.0514 arcsec	0.00318 arcsec

For Keck, we have $FM_a = 41.1\ FM_s$; for CELT, $FM_a = 952\ FM_s$. Note that the numerical coefficient is within a factor of order unity of the number of segments. The large magnitude of the CELT coefficient is a cause for concern. This is because there are anomalous fluctuations of focus mode, both predictable and otherwise, at Keck. As an example of the latter, every change of the sensor gain to the least sensitive range is accompanied by the introduction of a large amount of focus mode. This anomaly substantially complicates some Keck alignment tasks. This phenomenon could be caused, for example, by a small accompanying shift in some global voltage level. However, absent a detailed model, it is conservative to assume that the size of the effect in actuator space will be proportional to the above coefficient, or more than 20 times larger for CELT than it is for Keck.

Note that the uncompensated focus mode parameters are similar in Keck and CELT, but the compensated parameters are much smaller in the CELT case because of the larger number of smaller segments, an advantage of the CELT design.

8.6 System Overview (Telescope Control Wavefront Sensor)

As described above, CELT or CELT instruments will be equipped with enhanced guiders or telescope control wavefront sensors (TCWS) to control low spatial frequency degrees of freedom of the primary mirror or telescope aberrations as a whole. The TCWS is likely to resemble the active optics wavefront sensors of current large telescopes such as the VLT, Gemini, or Magellan. There is a high degree of commonality among the latter three wavefront sensors. Therefore in the interest of space, we will summarize the requirements, design, operations, and performance of only the VLT wavefront sensor. This choice does not reflect an overall value judgment, and we may eventually avail ourselves of design aspects of all three systems. The choice does, however, reflect the fact that at the time of this writing the VLT system appeared to be the best documented and most thoroughly tested of the three. Further information on the VLT system can be found in Guisard, Noethe, and Spyromilio (2000).

8.6.1 Example System Overview (VLT)

The primary mirror of the VLT is a thin meniscus mirror with a diameter of 8 m and a thickness of 175 mm. Continuous active control of the mirror figure is therefore necessary. Each of the four VLT telescopes has three foci, two Nasmyth foci as well as a Cassegrain focus; each focus has a permanently installed wavefront sensor.

The VLT Active Optics System (aOS) is a closed-loop system with corrections based on a modal concept. The corrected modes are Zernike defocus, third order coma, and the lowest elastic modes of the primary mirror. The corrections are applied by changing the position of the secondary mirror (defocus and coma) and the figure of the primary mirror (all other modes). The corrections are the result of measurements made using a Shack-Hartmann wavefront sensor.

8.6.2 Example System Guide Star Acquisition and Requirements

Continuing the VLT example, a guide star can be selected anywhere inside a field of view of 7.5 arcmin radius at the Cassegrain focus and 15 arcmin radius at the two Nasmyth foci. For comparison, science fields range from 2.5 to 4.5 arcmin in radius. The acquisition is done using a stage with two rotational degrees of freedom.

The frequency of corrections must be fast enough to compensate for gravitationally- or thermally-driven changes in the shape of the telescope structure, but slow enough to integrate over atmospheric-seeing-induced anisoplanicity effects. In practice, this requires integration times on the order of 30 seconds. The accuracy of the wavefront analysis is determined by the accuracy of the centroid determination on the Shack-Hartmann wavefront sensor. For bright guide stars, i.e., around magnitude 12, a centroiding error of 20 to 30 milli-arcsec is achieved. Simulations have shown that this accuracy is more than adequate for efficiently operating the system.

Fainter guide stars, down to a magnitude of approximately 14 (depending strongly on the color of the star), can also be used. If a star is too bright or too faint, the integration time of the exposure is automatically adjusted. For bright guide stars, the correction may be based on the average of wavefront analyses of several short exposures, so that a total integration time of 30 seconds is used. In principle, using guide stars fainter than magnitude 14 is possible by increasing the overall integration time. However, the error of the correction due to gravitational deformations of the primary mirror is larger than desired for integration times larger than approximately 60 seconds.

8.6.3 Optical and Mechanical Design of the Example System Wavefront Sensor

The VLT wavefront sensor is of Shack-Hartmann type using a lenslet array with 20 x 20 lenslets of diameter 500 microns and focal length 45 mm. The CCD has approximately 500 x 500 pixels with 20 pixels per subaperture. The readout time is 0.5 seconds. Water and a Peltier element provide cooling. The cost of the CCD was on the order of $50K. The criteria for the CCD selection were low read noise and, in particular, cosmetic excellence.

The design of the lenslet optics was chosen so that the spot size on the CCD is not smaller than 1.5 pixels. At low light levels the best results can be obtained with small spots (1.5 pixels). At high light levels (where the brightest pixels are close to saturation) the size can be larger. Tolerances for the optics are loose, as the WFS is a simple achromatic system.

The entire VLT aOS construction was subcontracted to Amos in Belgium and completed over a period of two years. Most individual components are standard, off-the-shelf items (notable exception: the CCD was custom made by Jena Optik). The optical and electrical designs as well as the conceptual mechanical design were done in-house at ESO. The software was also largely written at ESO. There were no major pacing items in the design.

The VLT provides a field derotator for the science instruments but no pupil derotator, as such a system is not necessary for a circular mirror telescope. The field derotator is not a part of the active optics WFS.

The lessons learned include:

- It would be desirable to have a second Shack-Hartmann sensor that measures only the relative alignment of the primary and secondary mirrors to correct for global astigmatism. With an integration time of 30 seconds, the residual atmospheric contribution to astigmatism is still too large to provide ideal results. (Note that the larger size of the CELT mirror will increase the size of the residual atmospheric effect [for a given integration time] because of the proportionally longer wind crossing time.) However, longer integration times should not be used, as the softest modes (elastic mode equivalent to third order astigmatism, defocus and decentering coma) change on time scales comparable to one minute.

- The aOS was designed to make a full wavefront analysis in 0.5 seconds. The expectation was that it would be used to correct for wind-induced perturbations using the primary mirror. Increases in computer speed are reducing image analysis time to about 1 second (0.5 seconds CCD readout and 0.5 seconds analysis).

8.6.4 Image Processing for the Example System Wavefront Sensor

- The frame is background subtracted.

- No flat-field correction is currently performed as this is not necessary to achieve the required performance.

- The centroiding algorithm for the Shack-Hartmann sensor uses a simple weighted average of pixels. Only pixels with counts above a certain threshold are used. Options for choosing the threshold are either 20% of the difference between the background and the maximum pixel, or 10 times the square root of the shot noise. A Gaussian fit to the subaperture images would only improve the centroiding accuracy by about 10-20%.

- No cosmic-ray protection is provided, as it would be difficult to implement and does not seem necessary. However, if only a single pixel in a subaperture is above a threshold, then this subaperture is ignored in the wavefront reconstruction.

- Wavefront reconstruction: After an interpolation of the transverse aberrations to a regular grid and an integration along rows and columns to obtain the wavefront error, the basis functions are fit directly. The integration also involves a least squares fit of the integration constants. The least squares fit and the fitting of the functions both use matrix multiplications.

8.6.5 Operation and Performance of the Example System Wavefront Sensor

After a preset to a new position in the sky, the images taken by the telescope are visibly deformed. Although the deformation is still small enough for the proper functioning of the algorithm for the wavefront analysis, the first correction is performed in open-loop based on a look-up table. The image quality after such an operation is approximately 1 arcsec FWHM. This could be improved, but since closed-loop operation starts immediately afterwards, an improvement is not necessary.

The system is extremely robust. Out of approximately 12,000 wavefront analyses per month per telescope, approximately three failures were noted. On average, two of those three were due to failures of the electronics of the CCD camera. The third is usually due to an inappropriate guide star, for example a double star.

The main residual error is due to residuals of atmospheric turbulence after integration over 30 seconds. The power in the modes is fit well by Kolmogorov turbulence.

All relevant data, in particular the coefficients of the modes and the residual variance, are logged for off-line processing.

The system is completely automated. It finds a guide star in a catalog and points to it automatically. As noted above, the first iteration is a coarse, open-loop correction. The first good correction is applied after the second iteration, that is, after slightly more than one minute.

8.6.6 Conclusions

Based on the experiences gained from the VLT aOS, the CELT TCWS will be feasible with current technology. The number and location of the CELT foci (there will be several) will have an effect on the design and cost of the TCWS. We need to decide whether to use several TCWSs, one at each focus, or to use a single system that is moved from one focus to another. In general, the TCWS will likely not be one of the critical components of CELT in terms of technical feasibility, schedule for design and construction, or cost.

8.7 Procedures for Optics Alignment

CELT optics will require measuring and setting more than 6500 degrees of freedom to define the positions and figures of the segmented primary mirror, the secondary, and the tertiary mirrors. We are currently developing a step-by-step procedure for systematically measuring and setting these degrees of freedom to an accuracy that meets the optical error budget. We will calculate the error propagation from measurements to image and wavefront quality. These will be described in a CELT report to be authored by Mast, Chanan, Nelson, and Noethe.

REFERENCES

Born, M. and E. Wolf. 1980. *Principles of Optics*. Pergamon Press.

Chanan, G., J. Nelson, T. Mast, P. Wizinowich, and B. Schaefer. 1994a."The W.M. Keck Telescope Phasing Camera System." Instrumentation in Astronomy VIII 2,198.

Chanan, G., T. Mast, J. Nelson, R. Cohen, and P. Wizinowich. 1994b. "Phasing the Mirror Segments of the W. M. Keck Telescope." SPIE Proceedings **2199** "Advanced Technology Optical Telescope V."

Chanan, G., M. Troy, F.G. Dekens, S. Michaels, J. Nelson, T. Mast, and D. Kirkman. 1998. "Phasing the Mirror Segments of the Keck Telescopes: the Broadband Phasing Algorithm." *Applied Optics* **37**,140-155.

Chanan, G., M. Troy, and E. Sirko. 1999. “Phase discontinuity sensing: a method for phasing segmented mirrors in the infrared.” *Applied Optics* **38**, 704.

Chanan, G., C. Ohara, and M. Troy. 2000. “Phasing the Mirror Segments of the Keck Telescopes II: the Narrowband Phasing Algorithm.” *Applied Optics* **39**, 4706-4714.

Guisard, S., L. Noethe, and J. Spyromilio. 2000. “Performance of active optics at the VLT.” SPIE Proceedings **4003**, 154-164.

Noll, R.J. 1976. “Zernike Polynomials and Atmospheric Turbulence.” *Journal of the Optical Society of America* **66**, 207-211.

Chapter 9. Adaptive Optics

9.1 Introduction

Large ground-based optical and infrared telescopes provide potential scientific benefits from both their increased light-gathering capability and their theoretically achievable image sharpness. To realize the advantage of resolution, however, it is necessary to overcome the serious degradation to image sharpness imposed by Earth's atmosphere. Even at the best sites in the world, under the very best "seeing" conditions, the effect of the atmosphere causes near-infrared images to be blurred to *forty times* their theoretical resolution for a 30-meter diameter telescope. Compensation of the blurring effects of atmospheric turbulence, through a technique known as adaptive optics, allows the recovery of the theoretical image sharpness.

Adaptive optics (AO) seeks to recover the theoretically possible performance of the telescope by correcting the deleterious effects of the Earth's atmosphere on wavefront coherence. By using a combination of sensors that measure the rapidly varying "imprint" of atmospheric index of refraction variations on incident light, and deformable mirrors that can apply an equal and opposite compensation, the most objectionable effect of atmospheric turbulence for ground-based imaging can be overcome.

Although astronomical adaptive optics was first proposed a half-century ago (Babcock 1953), only in the last decade of the 20th century have individual component technologies matured to allow the world's largest telescopes to enjoy adaptive optics correction. Today, while still a new addition to the astronomer's toolbag, AO systems are producing unique and valuable science (Wizinowich, et al., 2001) and a growing user community is bolstering the technology-led state of the art. Because the benefits of adaptive optics increase dramatically with increasing telescope diameter, AO has an essential and irreplaceable role in the design and use of the CELT observatory.

During the conceptual design phase, we have explored the technical feasibility of adaptive optics for a wide range of astronomical observations. In this chapter, we present error budgets and illustrative designs for two baseline AO capabilities, one optimized for mid-infrared and the other for near-infrared, that are intended to motivate the scale and nature of the AO development effort, without claim of system optimization. Detailed architectural decisions for first-light adaptive optics capabilities will be made during the next phase of CELT design.

9.1.1 The Science Benefits of Adaptive Optics

The impact of adaptive optics correction on the realizable science of CELT is enormous. In addition to the image resolution advantage described above, adaptive optics also provides improved imaging and spectroscopic sensitivity (growing as the square of the telescope diameter for background limited unresolved point sources, or as the telescope diameter D^4 for integration time); improved imaging contrast (when observing faint material in orbit around nearby stars); and increased spatial coherence (which allows efficient coupling to single mode fibers).

In addition, under seeing-limited conditions, the characteristic size of instrumentation used to parse and analyze the light collected by CELT will grow linearly as a function of telescope diameter, while the corresponding cost is proportionally higher. For imagery limited only by diffraction, the characteristic instrument size remains constant for any diameter telescope, resulting in cost savings when compared to instruments designed for only seeing-limited observations. Diffraction-limited images are physically smaller than seeing-limited images in the telescope focal surface. This can reduce the slit width for a

spectrograph and thus reduce the spectrograph size and cost, all other parameters (spectral resolution, field of view, etc.) being equal.

Two distinct adaptive optics instruments are envisioned to conduct the majority of the science programs discussed in Chapter 2. These instruments are known as low-order adaptive optics (LOAO), which will perform optimized low-background science at mid-IR wavelengths (3-30 μm); and multi-conjugate adaptive optics (MCAO), which is optimized for near-IR wavelength (1-2.5 μm) science. Specialized modes of AO operation, which may require dedicated instrumentation, can expand the available science beyond the two baseline AO capabilities. Two particular concepts are noteworthy: extremely high-contrast AO (EAO), and ground-layer AO (GLAO). The table below shows the AO capability required to carry out the representative science programs described in Chapter 2 (indicated by X) and the AO capabilities that would improve science results, but are of lower priority (indicated by *****).

Science Program	**LOAO**	**MCAO**	**EAO**	**GLAO**
Solar System science (Section 2.5.1)	X	X	*	
Spectroscopic search for terrestrial planets (Section 2.5.2)				
Direct imaging of terrestrial planets (Section 2.5.2)			X	*
Star formation (Section 2.5.3)	X	X	X	
Nearby galaxies (Section 2.5.4)		X		*
Galactic nuclei (Section 2.5.5)	X	X	*	
The z =1-5 universe (Section 2.5.6)		X		
Wide field science (Section 2.5.7)				*
The “dark ages” z > 5 (Section 2.5.8)	*	X		

9.1.2 Seeing-Limited Observations

The strong dependencies upon observing wavelength found in AO scaling laws leads one to conclude that fully correcting a 30-m diameter telescope at visible wavelengths over wide fields-of-view (FOV), though highly desirable, will be impractical on the time scale of CELT construction. Thus, the seeing-limited mode currently used on 8-10-m telescopes is likely to persist. Discussions among the CELT science community lead to estimates that roughly half of CELT science will require adaptive optics. The remaining science will be conducted in seeing-limited mode at visible wavelengths (typically over fields up to 20 arcmin in diameter).

9.1.3 Technical Development

The challenge of adaptive optics is among the greatest faced by CELT, requiring not only advances in component technologies (such as deformable mirrors, wavefront sensors, and guide star lasers) but also the development of new theory and techniques in order to minimize required technology investment. Fortunately, independent of CELT, adaptive optics technology will likely continue to mature (Gavel 2001), driven by the near-term desire to design and build AO systems that can correct the largest telescopes at visible science wavelengths (Dekany 2000). The increasing maturity of both the equipment and the community of users of these systems will help bridge the gap between the current generation of AO systems and those required for CELT.

9.1.4 Technology Scaling Laws

The technology requirements for any adaptive optics system are dominated by a set of scaling laws based upon the physical nature of turbulence in the Earth's atmosphere. In recent years, excellent books describing the systems engineering approach to AO development have been published (Hardy 1998; Roddier 1999). The most important scaling laws are summarized here for reference, assuming the science goal is to achieve a given Strehl ratio at observing wavelength, λ, on a telescope of diameter, D, under seeing conditions parameterized by r_0:

Required number of degrees of freedom ~ $D^2 / r_0^2 \sim D^2 / \lambda^{12/5}$
Required closed-loop correction bandwidth ~ $r_0^{-1} \sim \lambda^{-6/5}$
Required wavefront measurement photon flux ~ $r_0^{-3} \sim \lambda^{-18/5}$
Required level of control of systematic errors ~ λ

Derivation of these scaling laws can be found in Hardy (1998), showing that system requirements become steeply more difficult with increasing telescope diameter and decreasing observation wavelength.

9.1.5 Summary of AO Impact on Other Subsystems

The adaptive optics system requirements are related to many areas of observatory design requirements. The following list is a summary of important issues to be addressed, and constitutes the topics of focus in the preliminary design phase of CELT:

- The site must be selected to have favorable AO characteristics (seeing, cloud cover, and C_n^2 (h) distribution).

- Nodding, chopping, counternodding, counterchopping requirements may complicate the AO, telescope control, and instrument subsystems.

- Science requirements will drive both the LOAO architecture and the number and type of telescope foci.

- The pupil location and field curvature seen by the instruments is determined by details of the MCAO relay design.

- The primary mirror segment size and three-dimensional positional tolerance (in AO mode) is influenced by the number of actuators available for AO, and is also related to the limiting high-contrast observations achievable in EAO mode.

- The FOV and sampling density of instruments is determined by the image quality deliverable by AO over any given FOV.

- MCAO will benefit from real-time C_n^2 monitoring equipment. Queue scheduling may require meteorological prediction.

- Acceptable telescope windshake levels are coupled to the AO system's ability to reject tip/tilt.

- The acceptable emissivity budget of the telescope optics is determined by the mid-IR background seen by the LOAO capability

- Observatory safety systems and procedures (operations) must place particular emphasis on high-power laser safety

- The data archiving and pipelining process is coupled to the relevant AO data necessary for post-processing (such as image deconvolution).

- The required optical quality of the tertiary mirror is slightly increased due to the beam shear of the expanded diffraction-limited MCAO field of view (partial MCAO correction is likely over 2 arcmin FOV).

9.2 Assumptions and Definitions

9.2.1 Observing Assumptions

In order to facilitate quantitative analysis in the remainder of this chapter, we limit the potentially large parameter space for adaptive optics systems design. Specifically, we make the following assumptions regarding adaptive optics observations with CELT (see Hardy 1998 for definitions).

Wavelength of parameter definitions	λ_0	500	nm
Coherence cell size	r_0	20	cm
Greenwood frequency	f_G	50	Hz
Outer scale	L_0	> 30	m
Isoplanatic angle	θ_0	2	arcsec
Maximum off-zenith science angle	ζ_{max}	65	deg

(Although the telescope is expected to operate to zenith angles of 65 degrees, we assume 0 degree zenith angle for the calculations to follow.)

9.2.2 Turbulence

Vertical profiles

Seeing varies widely depending on site and season. We can make the assumption that one of the premier observing sites on earth will be selected, e.g., Mauna Kea in Hawaii; the high Andean desert of Chile; or one of a number of other sites that have low water vapor content, low atmospheric turbulence, and a high percentage of cloudless nights. Even among these sites there is considerable variability in the seeing parameters relevant to AO, possibly enough to affect decisions about the technologies to develop for AO. Site data from Mauna Kea (Roddier, Cowie, et al., 1990) suggests that much of the wavefront-distorting turbulence is spread over a range of high altitudes, 5-10 km above the summit, although the total integrated turbulence is among the lowest in the world. Cerro Pachon in Chile has higher integrated turbulence (worse seeing); however, it is concentrated in a ground layer and a single high altitude layer at 13 km, which makes the design of a wide-field AO corrector easier. It is clear that the choice of site, and the perception of the median turbulence profile for that site, will greatly affect the technology and design emphasis for the MCAO system. Site measurements of the C_n^2 profile will be important input to the final design of the AO systems. Furthermore, active monitoring of C_n^2 will be valuable in order to optimize the MCAO wavefront reconstruction on any particular observing night. Thus an on-line C_n^2 (h) monitor will be an important component of the system.

We will adopt the Kolmogorov model for the spatial frequency distribution of turbulence strength in the inertial range of frequencies (Hardy 1998). We use the standard definition of turbulence based on the refractive-index variance, C_n^2, and the integrated measures of turbulence of the coherent cell size, r_0, and isoplanatic angle, θ_0. Median C_n^2 profiles for Mauna Kea and Cerro Pachon (Ellerbroek and Riguat 2000) are shown in Figure9-1, along with a smooth fit to the Mauna Kea profile using a Hufnagel-Valley (Valley 1980) model having r_0 = 20 cm and θ_0 = 2 arcsec. Other authors (Roddier, et al., 1990) have estimated instances where θ_0 ~ 3 arcsec on Mauna Kea. Depending on the scheduling mode of the observatory, this larger value may in fact become the median value *during AO observations* (for a discussion of queue scheduling,see CELT Report No. 35, Nelson and Mast, ed., 2002). We adopt this model as the assumed profile for the remainder of this AO design chapter. This model is conservative in that the turbulence spread over a large altitude range may increase the requirements for MCAO in terms of required number of conjugate DMs and the number of laser guide stars.

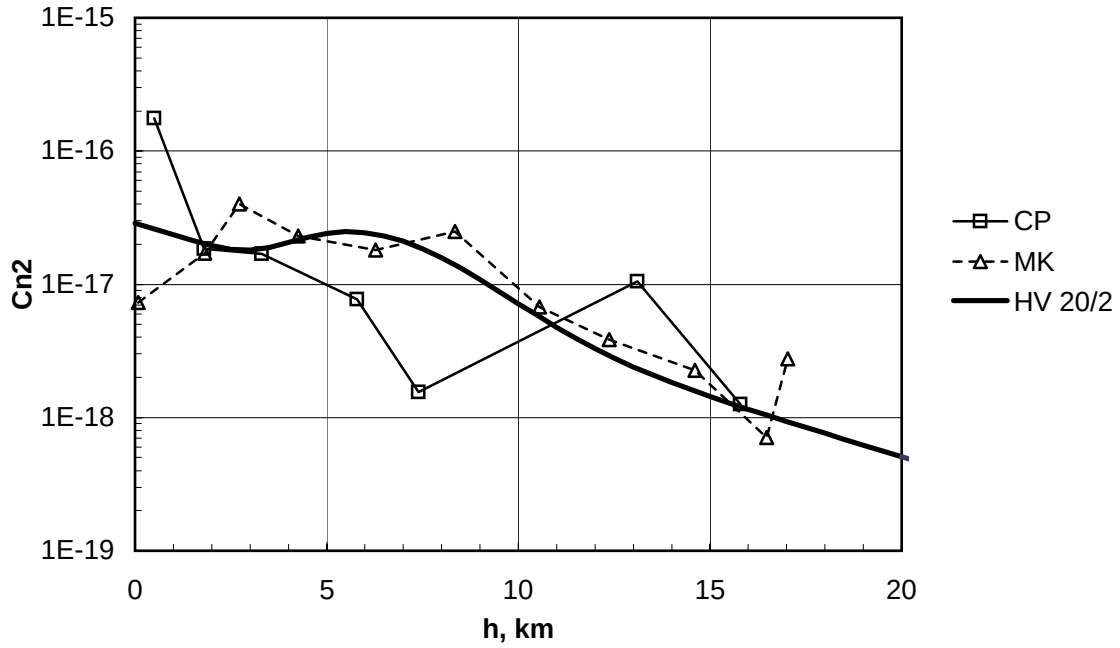


Figure 9-1. C_n^2 model profile. MK = Mauna Kea, CP = Cerro Pachon. Altitude is km above the observatory. The Hufnagel-Valley (HV) model is a smooth fit to the MK profile with r_0 = 20 cm and θ_0 = 2 arcsec.

9.2.3 Star Distribution Models

One of the key issues in adaptive optics is the availability of natural guide stars in any given field-of-view. An accepted model for the distribution of potential natural guide stars at visible wavelengths is the Bahcall-Soneira model (Bahcall and Soneira 1980), which gives the number of stars of a given magnitude per square degree at a given Galactic longitude and latitude on the sky. Since the AO system is likely to sense natural guide stars in the near infrared (1-2 μm) we adopt a similar distribution model for infrared wavelengths (Wainscoat 1992).

9.3 Low-Order Adaptive Optics (LOAO)

The strong wavelength dependencies in the adaptive optics scaling laws make diffraction-limited correction, over wide fields-of-view, relatively easy to achieve at mid-IR wavelengths, defined here as 3μm - 30μm. Fundamentally, this can be understood by noting that many effects in adaptive optics result in constant nanometers of wavefront error, representing a small fraction of these long wavelengths. The objectionable effect of atmospheric index of refraction fluctuations decreases rapidly for long-wavelength observations.

9.3.1 LOAO Error Budget

For the LOAO observing mode we adopt a performance goal of rms residual wavefront error of <500 nm over an FOV of 30 arcsec diameter. We allocate this error according to the following error budget:

Table 9-1. LOAO error budget

LOAO error term		**θ_0 = 3 arcsec nm, rms**		**θ_0 = 2 arcsec nm, rms**
Anisoplanatism (15″ to guide star)		225		319
Measurement	(m_K = 16.1)	270	(m_K = 14.8)	226
Controller bandwidth	(5.3 Hz)	238	(6.7 Hz)	195
Wavefront fitting	(400 actuators)	193		193
Non-correctable primary mirror figure		75		75
Non-correctable AO figure		40		40
Non-correctable instrument figure		40		40
Tilt measurement	(using higher order guide star)	106		70
Tilt bandwidth	(6.3 Hz)	74	(8.4 Hz)	49
Tilt anisoplanatism (using higher order guide star)		83		75
Total error		**500**		**500**

9.3.2 LOAO Sky Coverage

The primary technical advantage at mid-IR wavelengths is that the relatively large acceptable wavefront error allows the use of natural guide stars (NGS) over the entire sky. Using our model of infrared point source distribution (Wainscoat 1992), we find that the probability of locating an m_K = 14 magnitude guide star within a 30 arcsec FOV varies from 3% near the galactic pole (b = 90 deg) to 32% across a broad swath near the galactic equator (b = -10 to 10 deg for galactic longitudes l =-90 to 90 deg), to virtually 100% in the galactic equator. This fact, combined with the wide isoplanatic angle at mid-IR wavelengths (θ_0(3.5μm) = 32 arcsec and 21 arsec for our two model atmospheres), allows substantial sky coverage using only NGS.

LOAO illustrative design

The following point design approximately meets the above error budget.

Number of deformable mirrors	1
Number of actuators	380
Wavefront sensor (WFS) mean wavelength	H-band (1.65 μm)

WFS subaperture diameter	1.36 meters
WFS type	Pyramid or Shack-Hartmann
WFS integration time	18.9 msec / exposure
Controller –3db bandwidth	5.3 Hz
NGS brightness	84 photodetections / subaperture / exposure
NGS off-axis distance	15 arcsec
NGS magnitude (for high-order and T/T)	$m_K = 16.1$
Tip/Tilt observing bands	I, J, H, K
Tip/Tilt image diameter	Seeing-limited (~ 0.4 arcsec FWHM)
Tip/Tilt detector	IR quad cell (3e- read noise)
Tip/Tilt exposure time	15.8 msec / exposure

9.3.3 Emissivity Optimization Options

The implementation of even a low-order mid-IR AO system will be a significant advance over adaptive optics technology of today. For the purpose of evaluating the impact on the CELT observatory, we shall concentrate here on a discussion of the scientific advantage to be gained through careful minimization of the thermal background induced by the telescope and AO system combination. We consider here four options, ordered according to the thermal background they induce upon the mid-IR science observations:

Cryogenic AO system at prime focus	(2 warm surfaces)
AO system utilizing an adaptive secondary mirror or	
Cryogenic AO system at Nasmyth focus	(4 warm surfaces)
Non-emissivity-optimized AO system	(14 warm surfaces)

Dekany (2002) has calculated the relative performance of these options, and only the results will be presented here. First, we briefly describe each option.

Cryogenic AO system at prime focus

The ultimate low telescope background is achieved by performing mid-IR observations in a cryogenic environment at a prime focus observing station. By placing an entire adaptive optics system within a cooled cryostat, only one “warm” telescope reflection is seen (to be conservative, we also count the surfaces of the cryostat window as being warm). Cryogenic deformable mirrors today can allow operation down to 30 K, with good linearity and reasonable stroke (M. Ealey, Xinetics, Inc., private communication).

Normally, the optical field aberrations at prime focus are such that a large, expensive corrector is required. For CELT, however, we have determined that it is reasonable to cover a 1 arcmin science FOV using only a single elliptical reflector, when taking advantage of the highly segmented nature of the primary mirror. Nelson (2001) has shown the extraordinary result that, using the conceptual design segmentation scheme, the primary mirror is capable of reshaping itself from its slightly hyperbolic conic constant of -1.0027, through a parabola, to an elliptic constant of -0.9957, while retaining a residual (high spatial frequency) wavefront error of only 112 nm rms. This term is not included in the above error budget, but results in only a slight increase when combined in quadrature with the 500 nm rms allowed error.

A depiction of a prime focus adaptive optics relay (positioned atop the telescope), and a detailed optical design for this concept is shown in the following Figure 9-2. A pierced, flat deformable mirror is positioned just before the true prime focus, allowing the prime light to encounter the reimaging ellipsoid which forms a relayed image plane. The rms wavefront error from the optical design is < 200 nm over a 30 arcsec FOV and < 600 nm over a 1 arcmin FOV. The reimaged light is divided into a wavefront sensing arm that would likely consist of a pyramid wavefront sensor (Riccardi, et al., 1998), operating on natural guide stars and using near-infrared light. The science light would be delivered to an f/3 focal plane to feed one or more mid-IR cameras. Sky calibration, if necessary, may be accomplished by chopping the ellipsoidal mirror (assuming a reaction mass).

Unfortunately, the inclusion of a prime focus station has significant impact on the design of the telescope structure and enclosure. In addition to adding mass and reducing the structural stiffness of the top end of the telescope, the size of the telescope enclosure would likely increase from 42 meters inside radius to about 46 meters inside radius. This would significantly increase the enclosure cost.

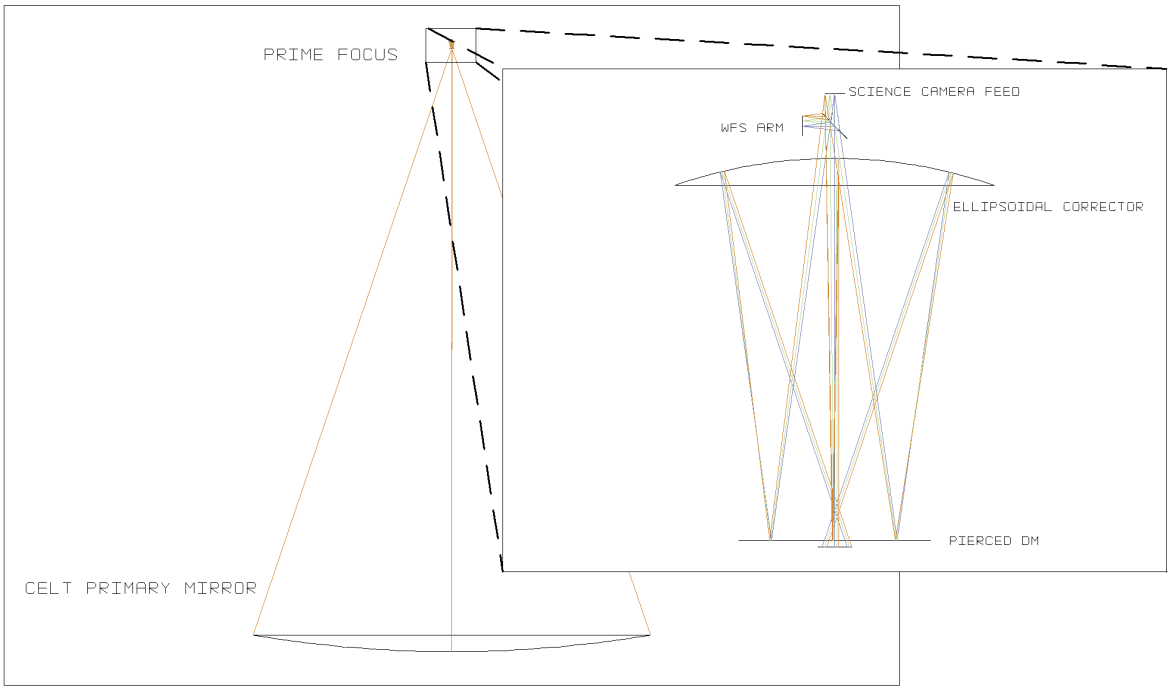


Figure 9-2. A prime focus, cryogenic adaptive optics system concept that provides good image quality over a 1 arcmin FOV, with an absolute minimum number of warm surfaces.

Adaptive secondary mirror

Another technique adopted for the Monolithic Mirror Telescope and Large Binocular Telescope (LBT) in southern Arizona (Lloyd-Hart, et al., 1998; Hill and Salinari 2000) is to reduce the total number of surfaces in a mid-IR system by making the secondary mirror of the telescope a deformable mirror. In these telescope designs, where a Cassegrain focus station is present, only two warm reflections are encountered before the light enters into a science camera cryostat.

For the case of CELT, however, we envision no Cassegrain focus, making adaptive secondary mirror technology less appealing since comparable background levels can be obtained with a cryogenically cooled AO system residing at Nasmyth (for a total of three warm telescope reflections).

Despite its thin facesheet, an adaptive secondary may require an increase in mass compared to a passive secondary mirror. Since the amount of mass supported at the secondary position is a major structural driver, increasing the secondary mass is currently thought to be a costly trade. Finally, adaptive secondary mirrors are not yet a mature technology. In particular, the diameter of the LBT secondary is only 1 m, whereas CELT requires a 4.2-m diameter secondary for a final focal ratio of f/15. Although we will continue to consider the option of using a 1-m class adaptive secondary in an f/60 configuration, the increased size of the field that must be accepted by Nasmyth instruments will likely complicate their design.

Cryogenic AO system at Nasmyth focus

Another alternative that provides as few warm reflections as the adaptive secondary option is the use of a cryogenic adaptive optics system at Nasmyth. The primary challenge of such a system is no longer the availability of effective cryogenic (30-40K) actuator technology (M. Ealey, Xinetics, Inc., private communication), but rather the system level difficulty of integration and test, and maintenance of a cryogenic system. At Nasmyth, a cryogenic AO system would suffer from background changes as the beam "walks" across the tertiary mirror. On the other hand, simplification could result from the stable gravity vector at Nasmyth. Although detailed study is needed, it is currently expected that the total observatory cost of this option would be less than either a prime focus cryo-AO system or an adaptive secondary system.

Non-emissivity-optimized AO system at Nasmyth

A final option, which carries the lowest cost, is to simply use a non-emissivity optimized AO system (i.e., the near-IR MCAO system) for mid-IR observations at Nasmyth. This would result in a large increase in the background, where we assume light would encounter some 11 additional warm surfaces, compared to a cryo-Nasmyth system. This would dramatically reduce the sensitivity of CELT at mid-IR wavelengths. Depending on the ultimate science priorities and detailed cost estimates of these alternatives, however, this may be an acceptable compromise as only the collecting area advantage of CELT would be lost, while the unprecedented imaging resolution would be maintained. As pure point source sensitivity at these wavelengths cannot compare to even small apertures in space, such as SIRTF, it is not clear what the ultimate loss of science under this option would be.

9.3.4 Emissivity Model Parameters

The system parameters for this comparison are shown in the following table.

Emissivity model assumptions

Air temperature	275K
Precipitable water vapor	3 mm
Primary mirror reflectivity	95%
Secondary mirror reflectivity	97%
Reflectivity of subsequent mirrors	98% each
Transmission of subsequent windows	98% each

Improvements to these reflectivities are worthwhile and may be possible. This will be studied in the next phase of work.

9.3.5 Relative Integration Times

The integration time required to achieve a given signal-to-noise ratio (S/N) on a sky background-limited object is the metric adopted by Dekany (2002). The fewer the number of warm optics (modeled as blackbody emitters) encountered by the mid-IR science beam, the shorter the required integration time. Because of variations in the emissivity of the Earth's atmosphere, the relative integration time is also a function of the observing wavelength. The relative integration time required for the four mid-IR options is shown in Figure 9-3.

Although large variations in the relative required integration time with wavelength can be seen, it is still useful to consider the average relative integration time as a measure of the scientific impact of these four options. To define the band edges, we take the average only over wavelengths where the atmospheric transmission exceeds 20%. These average integration time ratios are shown in the table below.

Table 9-2. Average ratio of integration time to achieve same S/N relative to cryogenic prime focus AO system

Wavelength Band	Cryo-Nas or adaptive secondary system	Warm MCAO system at Nasymth
K (2.2 μm)	1.32	3.31
L (3.4 μm)	1.51	4.76
M (5.0 μm)	1.34	3.51
N (10.6 μm)	1.35	5.41

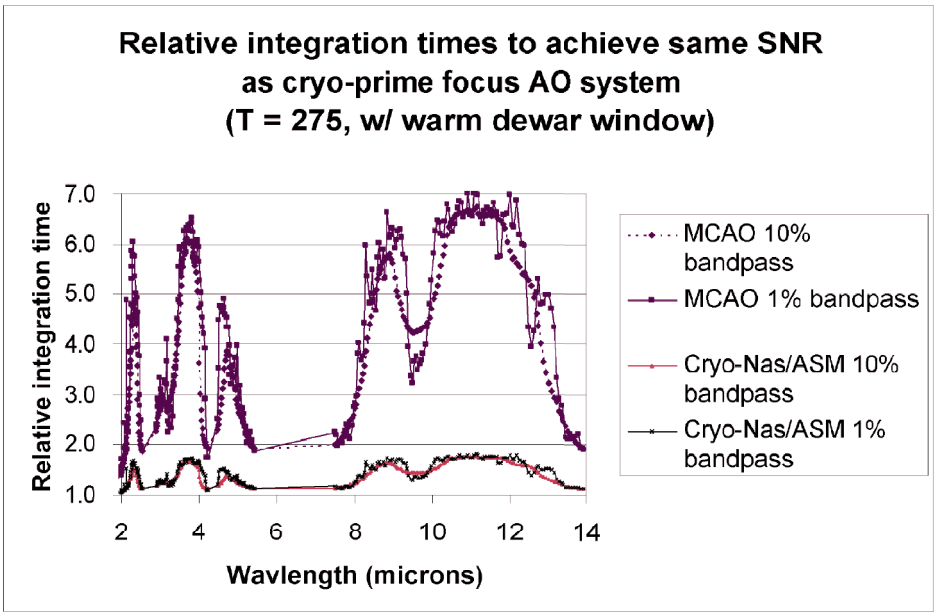


Figure 9-3. Relative integration time to detect a background-limited point source to the same S/N (for several emissivity optimization options) relative to the integration time for the ultimate optimization, a cryogenic prime focus system. The ratio is the same for both a cryo-Nasmyth system and an adaptive secondary, as CELT is not envisioned to have a Cassegrain focus.

9.3.6 L-Band

Observations at L-band, approximately 3.5 μm wavelength, represent a special case in the conceptual organization of AO observing modes presented herein. As seen from the discussion in Section 9.3.2, natural star density near the galactic pole does not support full coverage for L-band observations. One strategy could be to confine low-emissivity observations at L-band to the fraction of sky accessible via NGS, while allowing full-sky access at L-band through the MCAO system, albeit at significantly higher background.

In order to achieve full-sky emissivity-optimized L-band science, one may also employ the use of sodium laser guide stars. Because of the severe focal anisoplanatism of a single Na laser guide star, multiple LGS beacons will be needed, similar to the MCAO capability but without the need for multiple deformable mirrors. The presence of lasers at the telescope is very likely to be required for the near-IR AO capability described in Section 9.4 below. This would introduce additional complexity into any emissivity-optimized mid-IR AO system, but this is likely an acceptable alternative to conducting L-band observations through the high-emissivity near-IR MCAO system.

9.3.7 Outstanding LOAO Issues and Future Work

Given the quantitative comparison of relative sensitivity cost/benefit of each of the emissivity optimization strategies described above, the key decision to be made is the cost/benefit of each option, including consideration of the relative impact of each on the technical, schedule, and cost risk of the observatory.

Once a preliminary decision on the emissivity optimization strategy is made, work will likely concentrate on component technologies (such as cryogenic deformable mirror actuators) and preliminary systems requirements and design.

9.4 Multi-Conjugate Adaptive Optics (MCAO)

9.4.1 The Need for Multiple Laser Guide Stars

As the science observation wavelength decreases, the isoplanatic angle, θ_0, gets smaller (it scales as $\lambda^{-6/5}$) as does the coherence cell size r_0 and coherence time τ_0. Since, for a given quality of correction, the wavefront sensor's subaperture size must track r_0 and integration time must track τ_0, the required guide star brightness increases as $\lambda^{-18/5}$. Thus one needs brighter guide stars and a higher density of them on the sky. At observing wavelengths 3 μm and shorter, the probability of finding a sufficiently bright natural guide star within the isoplanatic patch of a dim science target becomes prohibitively low. Laser-generated artificial guide stars can be used in this case. Unfortunately, laser beacons can only be formed at a finite altitude (90 km for sodium beacons) and because the rays from the guide star travel to the telescope aperture in a cone, they do not probe all the atmosphere that affects the parallel rays coming from an astronomical object. This so-called cone effect is well known (Fried 1982) and forces the use of multiple laser guide stars for CELT. These stars are positioned at different field angles so that guide star rays probe the entire volume of turbulence that affects the science light. This concept can be extended to cover the volume of atmosphere that is used over a range of science field angles. For CELT, we want to cover a volume sufficient for a 1-2 arcmin diameter AO corrected field in the wavelength range of 1 to 3 μm.

9.4.2 Functional Requirements

At this point in the conceptual design of the CELT observatory, the functional and performance requirements for the MCAO system have not been finalized. The MCAO system architecture and the finalized set of requirements will be produced during the preliminary design phase. Our approach for addressing this important and complex issue is to adopt three levels of MCAO system performance: floor, requirement, and goal. The floor level of performance represents the minimal level of adaptive optics correction that meets the scientific goals of the observatory as a near-infrared diffraction-limited facility. The requirement level will be established as the target MCAO performance at "first lock" of the MCAO system on the sky. Finally, the goal level will be the expected performance that the MCAO system is expected to achieve within a decade of first lock, based upon on-going improvements in subsystem component technology. The MCAO system will be designed to support goal performance through the development of optomechanical, electrical, and software "backbones" that can be populated using this decade's technology (to meet the first lock performance requirements) and yet can be upgradeable to meet the goal performance levels using the technology of the next decade.

An example of possible performance floor, requirement, and goals levels of performance is shown in the following table, where RMS wavefront error is the worst case residual wavefront error within the specified field of view (FOV):

Table 9-3. Example MCAO system performance levels

RMS wavefront error (nm)	248	180	133
FOV (arcmin)	1	1	1
Sky coverage	30%	30%	50%
Strehl Ratio in J-band (1.25mm)	0.21	0.44	0.64
Strehl Ratio in H-band (1.65mm)	0.41	0.61	0.77
Strehl Ratio in K-band (2.2 mm)	0.61	0.77	0.87

In order to demonstrate feasibility of a 30-m telescope MCAO system meeting an aggressive goal level of performance, we performed a illustrative MCAO design during the conceptual design phase of CELT that achieves 133 nm RMS wavefront error over a 1 arcmin FOV, and 292 nm over a 2 arcmin FOV, over approximately 50% of the sky, limited by the availability of sufficiently bright natural guide stars.

Before describing the 133 nm system in detail, we note how some of the key technological requirements for the MCAO system vary depending on the target RMS wavefront error. Although detailed error budgets have been created for each level of performance, due to lack of space we highlight only a subset of the derived technological requirements:

Table 9-4. Example technology requirements

RMS wavefront error (nm)	248	180	133
Number of deformable mirrors	2	3	4
Total number of actuators	2,500	9,000	20,600
Number of sodium laser guide stars	5-7	7-9	9 + uplink AO

Meeting any of these performance levels will require investment in key technology areas, though we are confident that component technologies exist or can be straightforwardly developed to meet floor requirements. It is our intention, however, to seek and support technologies that can ultimately achieve the 133 nm rms wavefront error specification. As previously mentioned, determination of actual first lock requirements will be made during the preliminary design phase based on formalized CELT observatory science requirements and the demonstrated progress being achieved by industry, government, and academia in partnership with the CELT adaptive optics team.

9.4.3 Error Budget

A representative error budget for full MCAO correction is shown in the following table. These budgets are for the AO systems alone. The contributions from the telescope described in Section 11.3 must be added in quadrature.

Table 9-5. MCAO error budget

1 arcmin FOV Error Term	θ_0 = 3 arcsec nm, rms		θ_0 = 2 arcsec nm, rms
Generalized anisoplanatism	86		111
Wavefront fitting	60		60
Non-correctable primary mirror figure	50		50
Measurement	40		40
Tomography	32		37
Controller bandwidth, -3db (79 Hz)	25	(43 Hz)	42
Non-correctable internal AO figure	10		10
Non-correctable internal instrument figure	10		10
Tip/tilt bandwidth, -3db (50 Hz)	10	(70 Hz)	8
Residual tip/tilt anisoplanatism	26		26
Tip/tilt measurement (m_H = 15.3)	6	(m_H = 15.3)	6
Total RMS error	133		155
2 arcmin FOV Error Term	θ_0 = 3 arcsec nm, rms		θ_0 = 2 arcsec nm, rms
Generalized anisoplanatism	153		198
Other errors (same allocation as above)	102		109
Total RMS error	184		226

The error budget contains all of the usual AO terms. Fitting error is the error in approximating the atmospheric wavefront with a finite number of actuators on the deformable mirror. Controller bandwidth error is due to the control system's ability to correct wavefront variation to only a finite temporal frequency. Measurement noise is due to the Poisson statistics of the finite photocounts per correction cycle from guide stars plus any detector read noise.

Additional error terms are necessary to represent laser guide star MCAO. Tomography error arises from the finite number and placement of guide stars on the sky. Generalized anisoplanatism error results from the correction of the continuous atmosphere at only a finite number of conjugate layer altitudes.

For each of these cases, the error budget was optimized for the Hufnagel-Valley model atmosphere having r_0 = 20 cm and θ_0 = 3 arcsec that best fit the data from Ellerbroek and Rigaut (2000). The same system parameters would then be maintained and the error budgets for the corresponding r_0=20cm and θ_0 = 2 arcsec calculated, allowing for re-optimization of only the high-order controller and tip/tilt bandwidth terms.

In each of these error budgets we assume that the MCAO system is operating properly, and the residual tilt anisoplanatism is small compared to the case of NGS tilt anisoplanatism over a 2 arcmin FOV.

9.4.4 MCAO Illustrative Design

The following point design approximately meets the above error budget:

MCAO illustrative design

Number of deformable mirrors	4
Conjugate layer heights (above M1)	0, 3.0, 5.8, and 12.0 km
Number of actuators per DM (inscribed)	7700, 7000, 4200, 1780
Total number of DM actuators	20600
Wavefront sensor subaperture diameter	32 cm
Number of sodium laser guide stars	9
Brightness of laser guide stars	180 photodetections/subaperture/frame
Number of natural guide stars	fewer than nine, optimum unknown
Brightness of primary tip/tilt guide star	m_H = 15.1
Brightness of auxiliary tip/tilt guide stars	m_H < 23

MCAO relay optics issues

The design of any adaptive optics relay must balance competing scientific and technical goals. Fewer "warm" surfaces reduce emissivity and scatter, while more surfaces usually improve the imaging performance (particularly in large field designs). The design of an AO relay for a large telescope is much more difficult: It must consider all the prior challenges of AO relay design, and is additionally driven by the potentially large pupil demagnification. Moreover, the image plane location for an LGS formed at a given height moves according to the square of the ratio of telescope diameter (all else being equal). Thus, as the zenith angle (and hence distance to the sodium layer) changes, the LGS image plane moves by a much larger amount on larger telescopes. This can result in objectionable aberrations for the LGS light that depend on zenith angle, causing WFS calibration and system performance problems.

For the illustrative designs presented in this chapter, we have limited ourselves to flat DMs of diameter less than 35 cm. These DMs are considered relatively low-risk items on the time scale of CELT development. Some savings in number of surfaces are possible with optically powered (i.e., curved) DMs or larger DMs, and these are discussed herein.

The potentially significant aberration of LGSs requires special attention. Three solutions have been proposed to date:

- Use a dichroic near the Nasmyth focus to direct LGS light into a "trombone" that adjusts in length so that the LGS appears at the infinity focus position of the telescope. A second dichroic reinserts the LGS light into the AO relay. The trombone path length adjusts from 2.3 m (corresponding to the distance between the infinity focus and the 90 km LGS height at zenith) and about 1.5 m (corresponding to the distance between infinity focus and the LGS distance at a 45° zenith angle).

- Add passive optics or quasi-static active optics to the LGS part of the AO relay that "undo" the aberrations generated in the common path relay.

- Add an active optical element to each LGS WFS path that corrects the aberrations of the common path relay for that LGS as a function of zenith angle. These corrections would be known and could be calibrated. Since the light is monochromatic, phase-wrapping techniques are possible in addition to traditional continuous corrective surfaces.

These issues are addressed further later in this chapter and also by Bauman and Dekany (2002).

MCAO relay concept

The AO relay design for the MCAO configuration is particularly challenging because the relay should correct a 2 arcmin field and pass a 4 arcmin field without vignetting. (This leaves open the future option of exploiting the MCAO system in a GLAO mode.) Although the MCAO error budget calls for optimized DM locations at 0, 3.0, 5.8, and 12 km conjugate height, we shall in this section adopt DM conjugate heights 0, 4, 8, and 12 km in order to simplify the discussion of issues involved with the MCAO relay optical design, and to prove feasibility.

The layout for the MCAO relay is shown in Figure 9-4. The approach to avoid beam/optic interference described previously in the LOAO section is used again here: The DMs form "periscopes" that bring the light to a different beam height before it is folded back towards the center of the relay.

The performance of this relay is excellent, with an rms wavefront error of less than 35 nm throughout the entire 2 arcmin diameter field. The design is symmetric with f/15 in and out, and with exit pupil distance (with respect to the image) identical to that of the main telescope. The four off-axis parabolas (OAPs) are identical, each 3.3 m focal length, 750 mm diameter, and 570 mm off-axis.

Note that the image plane Petzval curvature currently is not corrected in the MCAO relay. Fixing the Petzval curvature in the MCAO relay would cost an extra warm surface. Rather, this can be accomplished in the science instrument with little S/N penalty since such surfaces are typically cold.

It is possible that the MCAO relay could be designed with fewer surfaces. However, the 4 arcmin field transmission requirement and 35 cm DM maximum mean that in a given "optical space" (i.e., the space between two powered optical surfaces), it is not possible to package DMs at arbitrary conjugate heights without mechanical interference. In fact, the closest that any two adjacent DMs can be is about 430 mm, which represents a conjugate height difference of 8 km. For example, DMs conjugate to 4 km and 12 km can be placed in the same optical space, but DMs conjugate to 4 km and 8 km cannot. Another

optical space, necessitating a powered optic, is required. For this illustrative design, with flat DMs at 0, 4, 8, and 12 km, it is not possible to meet all of the requirements with less than three powered optics. Due to packaging constraints, we have found the relay design to be much simplified if implemented with one additional powered optic, bringing the total number of OAPs to four. During the preliminary design phase, realizable designs with only three powered optics will be explored.

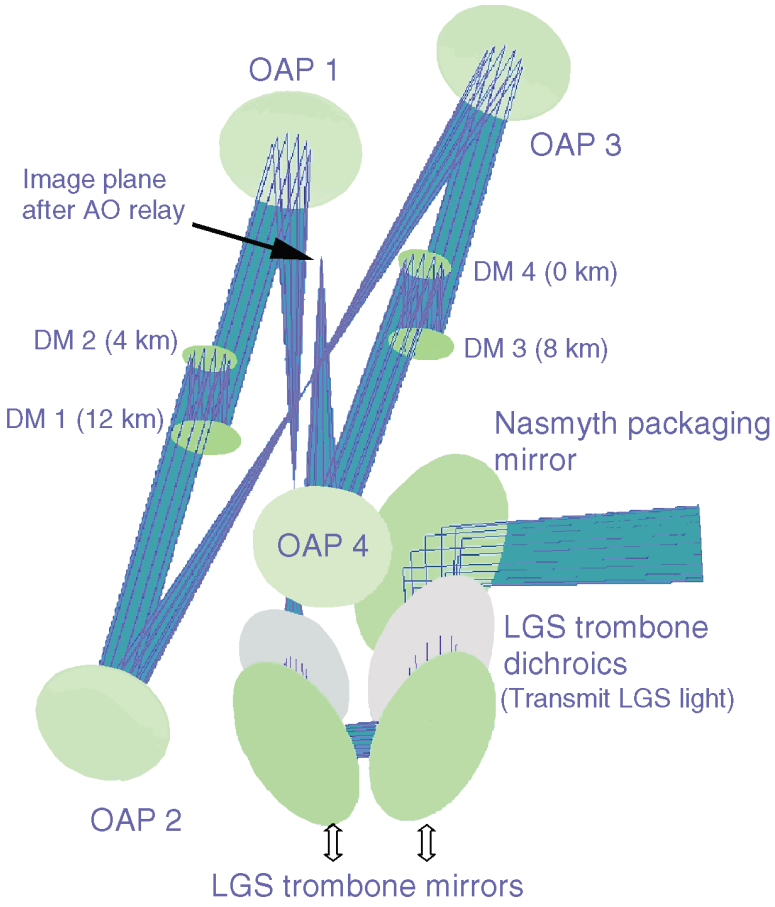


Figure 9-4. Optical layout of MCAO concept.

Reducing the required unvignetted field would also reduce the number of optics. If the required field were 2 arcmin, then DMs with conjugates 4 km apart could be placed in the same optical space. In the illustrative example, this would eliminate the need for a second optical space since all four DMs could be packaged together. Two surfaces would be eliminated, and the science path performance would improve since the OAPs could be made with longer focal lengths and smaller off-axis angles, yielding smaller aberrations.

Pupil and image rotation

Because CELT is an alt-az telescope and the MCAO system and instruments reside on the Nasmyth platform, both the image and pupil planes will rotate. Image rotation for the science path will be addressed through rigid body rotation of the instruments, nominally about a vertical axis, as discussed

in Chapter 10. Stopping down the scalloped pupil to the inscribed circle eliminates the dominant effect of pupil rotation in the wavefront sensor, namely the variation in the illumination fraction of perimeter subapertures. The light loss resulting from this is less than 3%, which is more than compensated by the absence of a K-mirror derotator in the science path (assuming internal pupil rotation for the instruments). The field rotation of multiple guide star asterisms does need to be compensated in the wavefront sensor. There are several approaches to this, including articulating pairs of mirrors for each guide star or a rotating wheel of pyramids, depending on the wavefront sensor architecture adopted after preliminary design. Boresight errors can appear as global tilt errors in the science path, but on-instrument wavefront sensors (OIWFS) would nominally detect this.

Wavefront sensing requirements

Shack-Hartmann wavefront sensors provide a high quantum efficiency solution to sensing wavefronts at the sodium guide star wavelength of 589 nm. Since the on-axis wavefront solution will not necessarily flatten the wavefront from each guide star, it is envisioned that the WFSs will operate in a non-null-seeking mode; hence there is a requirement that the sensor produce a linear off-null measurement of subaperture wavefront slope. Assuming that 4 x 4 pixels per subaperture are used on 32 cm subapertures, on the order of 375 pixels across the detector are required. The wavefront controller bandwidth must achieve 79 Hz -3db closed-loop bandwidth, demanding 1.6 kHz frame rate (0.625 ms per exposure) assuming conservative control laws. In addition to operating rapidly, such detectors must also have read noise less than the typical photon shot noise (~13 photons rms). Although such detectors are not currently available, they are within a few years technology development, given sufficient investment levels, based upon the historical rate of progress in low noise detector development.

The detector size and modest laser power requirement (see below) assume a scheme for reducing the LGS elongation that occurs when viewing the 10 km thick mesospheric sodium layer for an off-axis subaperture. This correction can be achieved through either a dynamic refocusing element or by synchronized WFS CCD clocking, assuming a custom radial geometry design of pixel locations on the CCD. Discussions with CCD experts (J. Beletic, private communication) support the feasibility of such custom detector arrays.

Deformable mirror requirements

Because of the baseline desire for wide field-of-view correction, multiple deformable mirrors must be used in order to separately correct for turbulent layers at various altitudes in the atmosphere. Since the atmospheric turbulence is not necessarily distributed in defined layers, and heights of the dominant turbulence are not necessarily known or constant, there is an anisoplanatism error corresponding to the placement of the DMs at conjugate altitude locations within the AO system. Tokovinin and LeLouarn (2000) have determined a method of characterizing this error, which they named “generalized anisoplanatism.” Figure 9-5 shows generalized anisoplanatism as a function of the number of DMs, each of which is placed at an optimal conjugate heights for a given C_n^2 profile.

In order to meet the error budget allocation for generalized anisoplanatism, a total of four DMs are required, optically conjugated to heights 0, 3.0, 5.8, and 12.0 km above the telescope. For the purposes of illustration, we can calculate the number of actuators required on each of the deformable mirrors in order to meet the fitting error allocation within the error budget. The result is that approximately 7700, 7000, 4200, and 1780 actuators, respectively, are needed on the four mirrors (inscribed in the projected 2 arcmin FOV metapupil at the conjugate height of each DM, assuming no central obscuration).

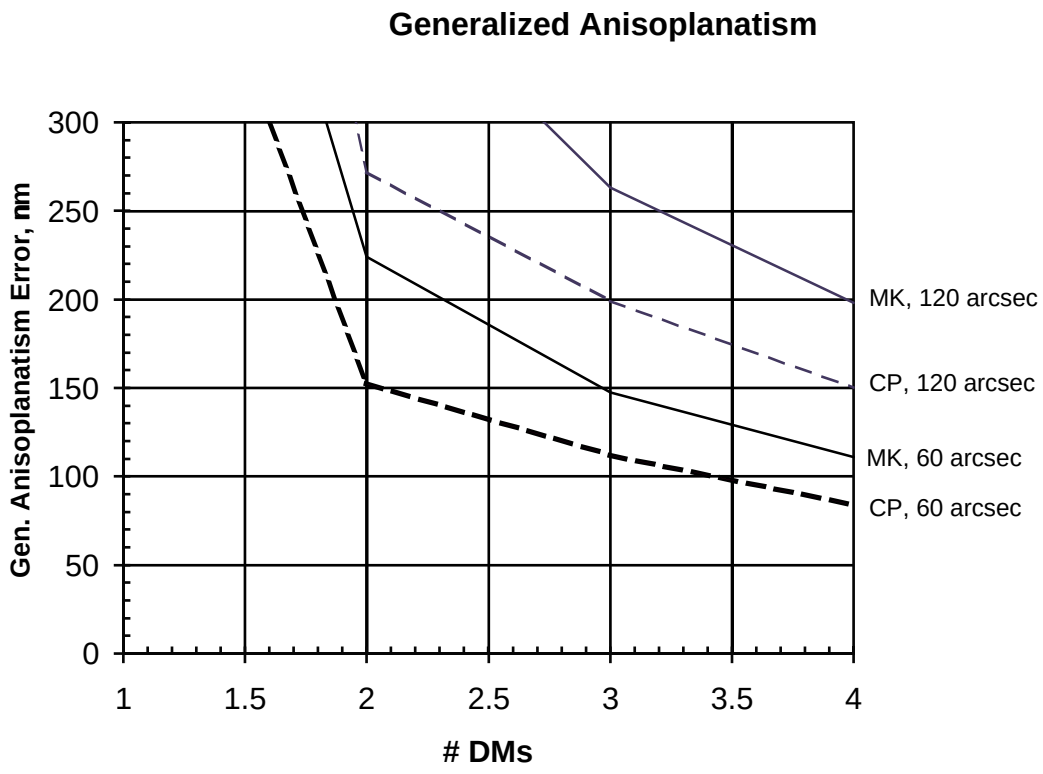


Figure 9-5. Generalized Anisoplanatism vs. number of DMs, for both a Mauna Kea (MK) and a Cerro Pachon (CP) atmospheric models, for two fields of view, 60 and 120 arcsec diameter.

Laser guide star requirements

The laser guide stars must be appropriately positioned and sufficiently bright to probe the expected turbulence profile in the atmosphere. Tomographic error is defined as the sampling error due to the finite number of guide stars and their geometric configuration on the sky. We utilize a method developed by Tokovinin and Viard (2001) to estimate it. For purposes of illustration, we imagine a circular pattern of guide stars which would need to be 15 m / 90 km = 34 arcsec in radius to probe the cylinder of atmosphere above the telescope. Figure 9-6 shows the error as a function of number of guide stars and guide star constellation radius.

The brightness of the guide stars determines the measurement error component of the error budget. Figure 9-7 shows the measurement error as a function of guide star signal. We have assumed that the wavefront sensor is a Shack-Hartmann sensor with 32 cm subapertures, and that the LGS spot in the sky is 0.4 arcsec full-width at half-maximum. Scaling from typical, measured Lick Observatory sodium guide star return numbers, 180 photodetections/subaperture/frame (at a frame rate of 1.6 kHz) can be expected from an 11 watt laser beacon. To enable this relatively low total laser power requirement, it is important to keep the guide star extent on the sky small. Thus we envision the AO correction of the beam at the launch telescope to remove atmospheric aberrations in the uplink path. Conversely, if it proves more economical, we could accept the atmosphere-limited 1 arcsec laser spot, but at the cost of $(1/0.4)^2 = 6.25$ times more laser power per beacon.

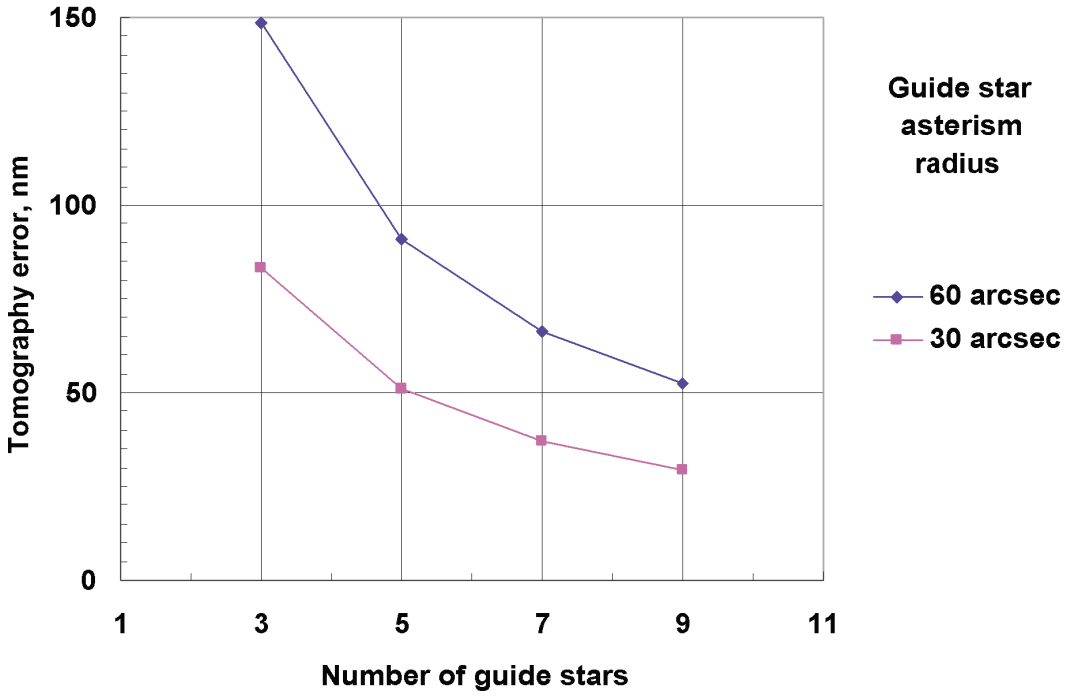


Figure 9-6. Tomographic error vs. number of laser guide stars.

It is difficult to disentangle the effect of measurement error (caused by photon noise) and tomographic error (caused by imperfect sampling of the atmospheric volume above the telescope) . Tokovinin and Viard (2001) estimate the two effects together, along with the noise-free case. In order to estimate the measurement noise contribution, we recreated their analysis, and subtracted the tomographic error terms in quadrature from the combined error, for a number of different brightnesses and number of guide stars.

The resultant curves shown in Figure 9-7 demonstrate that for any given asterism, increasing photoflux per guide star reduces the measurement noise term. We can fit an empirical function to the measurement error contribution near the optimum with

$$\sigma_{meas}(nm) = 839\ n_{gs}^{-0.39}\ (\alpha^2/n_{ph})^{0.314} \qquad (9\text{-}1)$$

where n_{gs} is the number of guide stars, n_{ph} is the number of detected photons per subap per exposure per guide star, and α is the LGS spot size in arcseconds. If we further define $N_{ph} = n_{gs} * n_{ph}$ = the total number of detected photons per subap per exposure for all guide stars, we find the interesting result that

$$\sigma_{meas}(nm) = 839\ N_{ph}^{-0.39}\ \alpha^{0.628}\ n_{gs}^{-0.076} \qquad (9\text{-}2)$$

Thus, the important quantities to reducing measurement noise is the total photoflux (which scales with broadcast laser power) and the guide star spot size, not the distribution of the laser photons into greater or smaller numbers of guide stars, n_{gs}. There is a small, but nonzero, reduction of the measurement noise term when increasing the number of guide stars, keeping the total photoflux constant.

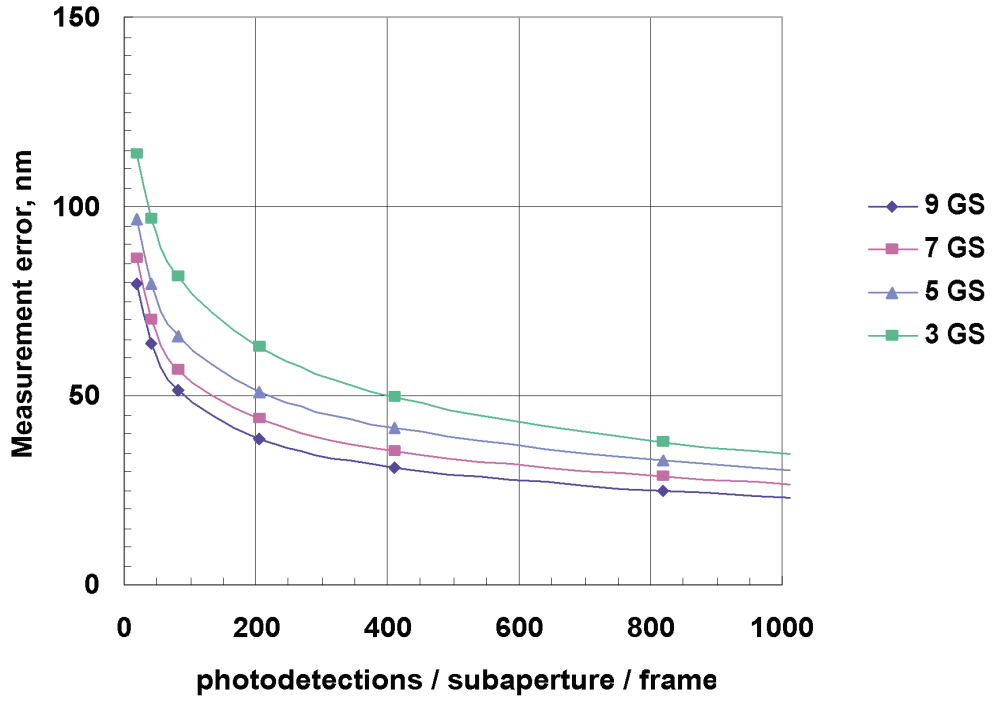


Figure 9-7. Measurement error vs. laser return per guide star, assuming the LGS spot size = 0.4 arcsec FWHM.

To adjust for the science field of view, it is desirable to be able to steer the guide stars to an optimal pattern. For example, as seen in Figure 9-6, the tomographic error is reduced inside a smaller field of view if the diameter of the guide star constellation is similarly reduced.

The finite extent of the Earth's sodium layer results in an elongation of the sodium guide star when viewed from subapertures displaced from the launch telescope. The elongation in radians, $\varepsilon = \cos z \; L*x/h^2$ where L is the excited sodium length (e.g., the sodium layer thickness), x is the subaperture distance from the launch telescope, h is the layer height, and z is the zenith angle. For a subaperture with x = 15 m, the elongation is approximately 3.8 arcsec.

The schemes under consideration for removing guide star elongation include dynamically tracking a laser pulse as it traverses the sodium layer. This will require that the laser be pulsed so that only one pulse is in the sodium layer at a time. Furthermore, the pulse duration must be short enough that the spot appears smaller than the required 0.4 arcsec for wavefront sensing accuracy. Considering the speed of light through the 10 km thick sodium layer, this works out to be < 2.0 microsec pulse duration. The maximum pulse repetition frequency is 7.5 kHz at z = 60 degree zenith angle, and 15.2 kHz at z = 0 degrees. Implementation possibilities include very fast imaging, synchronous change transfer on special CCDs, and real-time mechanical refocusing of the pulse traversing the sodium layer.

An alternative scheme to overcome the problem of sodium laser guide star elongation is to use the fact that for a given subaperture, one can get multiple spots, elongated in different directions (corresponding to the different laser launch telescope positions). One may then be able to achieve the same quality of wavefront measurement by making roughly twice as many spots and relying on the narrow direction of

each spot. This requires twice the laser power as would be needed for guide stars that are not elongated, but avoids the costly penalty of increasing laser power as the *square* of the elongation ratio. For example, compared to a 1 arcsec LGS spot, a 3.8 arcsec elongated spot would require > 10 times the laser power to reach the same wavefront slope measurement S/N. This scheme does not require any special pulse format.

Natural guide star requirements

Additional natural guide stars, other than a single tip/tilt reference, are needed for MCAO. Since the laser guide stars traverse the atmosphere on the upgoing path, their positions on the sky are not fixed a priori with respect to the background stars. As a result, distortions of the focal plane are introduced because the LGS wavefront sensors are insensitive to absolute tilts of the guide stars. A number of natural guide stars are needed to sense the distortion modes. These can be much dimmer than the wavefront reference stars since the whole telescope aperture can be used to recover just the tip and tilt components of their wavefronts. The number of natural guide stars required is not yet known, but if they are located within an isokinetic angle of the laser guide stars (about 2 arcmin), then we can expect that, as an upper limit, no more natural guide stars are needed than laser guide stars. One issue is the probability of finding enough suitably bright natural guide stars within the isokinetic angle. Basic calculations using visible (Bahcall and Soneira 1980) and near-infrared models (Robin and Creze 1986) show that if the natural guide stars are sensed in the infrared (to take advantage of the AO correction), then very dim guide stars (~23 magnitude) can be used. According to the star distribution models, there are significant numbers of these to give reasonable sky coverage for MCAO.

Global tip/tilt information can be acquired from the brightest NGS in the MCAO field of view. To enable 50% sky coverage, we may use a NGS of $m_H = 15.3$ with the 2 arcmin FOV, which satisfies the error budget performance goals, assuming a tip/tilt star integration time of 2.0 msec, and a -3db closed-loop bandwidth of 50 Hz. Determining the efficacy of multiple faint NGS correction of tilt anisoplanatism requires detailed modeling codes that we could not develop within the resource constraints of the conceptual design phase. Should adequate tilt anisoplanatism suppression require NGS significantly brighter than $m_H = 23\text{-}25$, we can alternately solve for tilt anisoplanatism using Rayleigh beacons in conjunction with sodium beacons in the hybrid scheme suggested by Ellerbroek and Rigaut (2000). We intend to explore the performance, complexity, and cost tradeoffs between the NGS and Rayleigh LGS solutions to tilt anisoplanatism in detail during the preliminary design phase.

Computing requirements

Exact computational requirements remain unknown, depending on the technological advancement of fast wavefront reconstruction algorithms coupled with improvements in computer hardware architecture and speed. Simply scaling today's algorithms, the computing requirement could be as great as a 20k x 40k vector-matrix-multiply (8×10^8 operations) per 0.625 milliseconds. Compared to the present day Keck AO system operations of $349 \times 640 = 2.2 \times 10^5$ ops per 1.6 ms, this is about a factor of 4,000 greater. Advanced sparse matrix techniques, as well as Fourier transform techniques, show promise of significantly reducing the calculations. We have demonstrated the significant advantage of sparse matrix methods at Palomar (F. Shi, private communication) and are confident that these techniques can be extended for CELT application. We can hypothesize that sparse algorithms will result in computations of order $2 \times 80k \times \log_2(80k) = 2.6 \times 10^5$ operations per 0.625 msec, about a factor of 30 greater than the present computer load for Keck AO.

9.4.5 Focal Anisoplanatism (FA) Mode (on-axis correction only)

Adaptive optics correction of CELT over a field of view limited by the isoplanatic angle only requires a single deformable mirror; the goal is to correct only the on-axis wavefront to high Strehl ratio. Multiple laser beacons and their associated wavefront sensors are still needed to correctly measure all of the turbulence that is in the path of the incoming starlight -- that is, to overcome the cone-effect of a single artificial beacon. We call this mode of correction "FA mode" to indicate the removal of focal anisoplanatism. It is instructive to compare the complexity, cost, and science tradeoffs between the FA mode and the full MCAO mode described in Section 9.4.4.

FA mode error budget

FA mode does not have the generalized anisoplanatism term since observing is on-axis. We add a term in the error budget for cone-effect error (focal anisoplanatism), which is the error on-axis due to residual under-sampling of the atmosphere by finite altitude artificial guide stars. An arbitrary but hopefully reasonable amount for this error is assigned for now, since the state-of-the-art for modeling tomographic error in a multiple guide star geometry assumes beacons are at infinite distance. Methods for quantifying the cone effect in an MCAO system are under investigation.

The following table shows the error budget for on-axis diffraction-limited AO (FA mode).

FA mode error term	**nm, rms**
Wavefront fitting	70
Focal anisoplanatism	10
Tomography	34
High-order Measurement	40
Controller Bandwidth	42
Tilt anisoplanatism	66
Tilt measurement	79
Tilt bandwidth	79
Total error	**162**

The following point design approximately meets the FA mode error budget.

FA mode illustrative design

Number of deformable mirrors	1
Number of actuators	5,000
Wavefront sensor subaperture diameter	38 cm
Controller -3db bandwidth	120 Hz
Number of sodium laser guide stars	9
Brightness of guide stars	180 photodetections/subaperture/frame
Tip/tilt guide star distance	20 arsec
Brightness of tip/tilt guide star	mK = 20
Tip/tilt -3db bandwidth	5 Hz

9.4.6 Outstanding MCAO Issues and Future Work

The following areas of technological development have been identified as critical for MCAO on CELT. Additional discussion of large telescope AO technology challenges can be found in Gavel (2001).

High actuator count deformable mirrors
A key feature of MCAO is the need to drive a number of deformable mirrors to correct for atmospheric turbulence at multiple conjugate altitudes. The high number of actuators (~21,000) at the present cost per actuator ($500/actuator) makes using existing DMs cost-prohibitive. The development of micromechanical DMs that might cost about $10-$20 per actuator, or integrated electrostrictive actuators at $50 per channel are both attractive alternatives to be pursued in the next phase of CELT.

Large format high speed CCDs
CELT MCAO calls for detectors with on the order of 400 x 400 pixels, capable of reading out at 1200 frames per second with < 5 electrons read noise. Current state-of-the-art has achieved this frame rate and read noise, but not this size (128 x 128 is the largest to date). In addition, the need to remove laser guide star elongation effects may require a CELT-specific pixel geometry and non-traditional clocking, and would require the design of a specialized CCD chip.

Fast wavefront reconstruction algorithms
MCAO will also have significant real-time computational requirements. Consequently, developing a computationally tractable algorithm for this purpose is an important technology milestone.

High power sodium lasers
CELT MCAO will require several high-power (~10 watt) lasers at 589 nm wavelength, possibly with a specialized pulse format. Technological development of such lasers is still in its infancy, and is currently being pursued by the Gemini telescope MCAO project (Gemini Observatory 2001) and by the Center for Adaptive Optics located at University of California Santa Cruz. Presently, solid state lasers of about 2 watts and dye lasers of up to 15 watts (pulsed) have been used in astronomical AO systems. Substantial effort is needed to produce efficient and robust designs for use at CELT.

Open-loop wavefront sensing
MCAO systems for 30-m class telescopes are likely to require a significant increase in systematic error control and calibration over existing adaptive optics systems. Presently envisaged tomography architectures require wavefront sensors to operate partially or entirely in “off-null.” Since the wavefront sensor will then need to operate linearly over a significant dynamic range, the null-seeking advantage enjoyed by all existing wavefront sensors will likely be lost. Deformable mirrors will also need to be accurately calibrated since they will influence the sensor readings. More attention to calibration and metrology of the entire AO system will be required to maintain high wavefront quality.

Systems prototyping
The complexity of CELT MCAO, measured by the number of new components, subsystems, and system-level technologies, is likely to require significant prototyping to determine practical limitations to real-world implementation. Notable areas that cannot be sufficiently addressed through simulation are: hardware-specific implementation of wavefront reconstruction algorithms; non-common path wavefront calibration between the MCAO wavefront sensors and science instrumentation; and practical issues of LGS operation, which remains in its infancy for even single guide star, single conjugate AO systems.

9.4.7 Impact of MCAO on Other CELT Subsystems

By design, the MCAO concept presented here has minimal impact on the telescope structure and optical design. It is unlikely that an adaptive secondary mirror meeting the required specifications for MCAO (4.0-m diameter with > 6000 degrees of freedom) will be economically practical in the coming decade. The size of the enclosure opening for CELT is larger than the diameter of the primary mirror, due to the need of placing laser guide star launch telescopes around the periphery of the telescope. Packaging of the MCAO and diffraction-limited science instruments may drive the size and structural performance of one of the two Nasmyth platforms. Because of this uncertainty, we have assumed that each AO science instrument will provide fast tip/tilt and focus information via an OIWFS, as discussed in Chapter 10. Of the most significant cost impacts will be the software required to integrate with the instruments and telescope control software, a very complex MCAO system containing: multiple lasers, WFS, and DMs; a sophisticated acquisition system; unprecedented calibration requirements, nodding/ counternodding; ADC control; safety systems; and etc. A comprehensive software architecture for the entire observatory must be developed with the goal of substantially reducing operations costs. This was not the software development process followed for Keck Observatory.

9.4.8 Impact of Non-Favorable C_n^2 Profiles

The C_n^2 profile and its variability could have significant impact on the architecture of the MCAO system. Turbulence spread evenly over a wide range of altitudes is most difficult for MCAO, requiring more DMs and laser guide stars. On the other hand, although turbulence may be concentrated in a few sharp layers, worse overall seeing (small r_0) is unfavorable in terms of increased number of actuators, smaller subapertures, and consequently more laser power. Both C_n^2 magnitude and vertical profile information should be taken into account in the final site selection. Site surveys should plan on collecting this data with sampling frequency of a few hours or less, over baselines of years.

9.4.9 Alternative MCAO Architectures

Architecture of adaptive optics systems for 30 m class telescopes is a rapidly evolving area of research. In this section, we present one illustrative architecture designed to convey the feasibility of MCAO for CELT. Numerous alternative wavefront sensing and wavefront correction approaches have been suggested in the literature, each with its own areas of promise and challenge. During the next phase of the CELT project, a more careful trade study of these approaches (and, undoubtedly, creative new ones) must be undertaken to minimize the cost of MCAO.

Several alternative architectures to our illustrative design are presented in Table 9-6.

9.5 Additional Science AO Observing Modes

9.5.1 Extreme Contrast Adaptive Optics (EAO)

Introduction and goals

One of the most exciting science applications for CELT is the direct detection, through reflected light, of extrasolar planetary systems. Although the detection of certain very young extrasolar planets (Section 2.5.2) may be possible with the baseline MCAO system (Section 9.4), we consider here the more challenging problem of detecting mature extrasolar planetary systems resembling our own.

Table 9-6. MCAO Architecture Concepts

Architecture concept	Features	Advantages	Disadvantages
CELT conceptual design presented here	Multiple Na LGS, multiple NGS, using Shack-Hartmann WFSs.	Architecture is feasible extrapolation from existing LGS AO systems. Tomographic advantage of multiple guide stars.	Requires development of Na guide star lasers and Na LGS observing techniques. Requires aircraft and satellite avoidance.
Layer-oriented wavefront sensing (Raggazonni 2000)	Pyramid WFS, possibly using NGS only for partial sky coverage.	One WFS camera per deformable mirror, S/N f gain for near-IR NGS.	Optical design difficulties or ELTs.
Stitching (i.e., Tyler 1994; Fried 1995)	Multiple Rayleigh LGS, NGS to overcome poorly sensed low order spatial modes.	Uses less sophisticated laser technology.	Increased system complexity and computer requirements.
Plasma guide stars (Ribak 1998)	High powered radio transmitters used to generate airglow guide stars.	Avoids use of sophisticated lasers.	Requires MW of transmitted power.
Curvature system (Roddier 1991)	Curvature WFS using NGS only.	Avoids use of LGS, reduced actuator count.	Limited sky coverage and correction only down to K-, possibly H-band.

A mature planet such as Jupiter is approximately 10^9 fainter than its parent star -- far beyond the reach of current AO or space telescopes. Angel (1994) recognized that an AO system with 10^4-10^5 degrees of freedom on 6.5-10-m telescopes could approach these contrast levels and detect extrasolar planets around a handful of the nearest stars. Later papers (Macintosh, et al., 2001) have begun to explore the scaling laws governing these systems in more detail. The sensitivity of such systems increases very rapidly with telescope size, making them much more compelling for a 30-m telescope than an 8-m. Although the technological requirements are daunting, recent developments in areas such as micro-electro-mechanical (MEMS) deformable mirrors and high speed computers are bringing these systems within the realm of possibility.

Such systems have now been dubbed extreme contrast adaptive optics, or EAO. These can be defined as AO systems whose primary concern is not improving the central peak, but suppressing the halo of scatted light that surrounds any AO-corrected star. EAO systems have a characteristic PSF with a sharp peak and a flat, extremely faint halo. Figure 9-8 shows the contrast ratio -- defined as the minimum brightness ratio of a detectable point source to the brightness of the central star -- as a function of radius, for two target stars. A CELT EAO system should easily be able to direct image Jupiter analogues around nearby stars, and can even -- if it achieves its theoretical maximum performance -- marginally detect Earth-sized planets around the very nearest stars.

EAO performance analysis

Traditional AO system error budgets, such as those in preceding sections, emphasize the Strehl ratio. The relevant parameter for EAO systems, as discussed above, is really the intensity of the residual halo and the noise, or deviations from smoothness, of this halo. Angel (1994) shows that the intensity of the

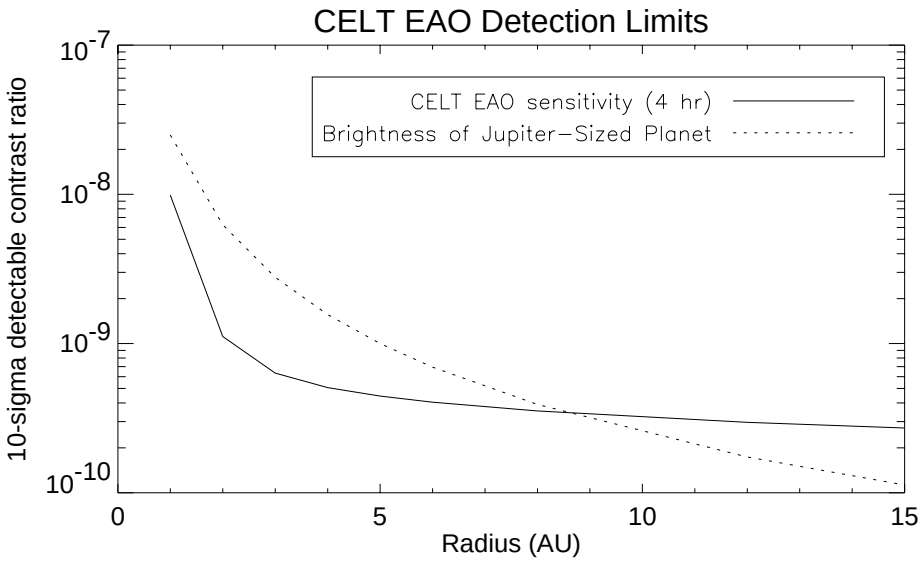


Figure 9-8. Detectable contrast ratio between planet and star as a function of separation for a target star at 4 pc and the EAO system described below, for a four-hour integration. Overlaid is the brightness of a Jupiter-like planet in the target system.

halo normalized to the peak intensity of the star (which he calls the gain, G) is given by $G=(1/\sigma^2)(D^2/d^2)$ where D = telescope diameter, d = subaperture size, and σ^2 is the total wavefront variance. In the idealcase, σ^2 can be broken down into three irreducible terms, which we will designate σ^2_{wfs}, σ^2_{fit}, σ^2_{bw},. These terms correspond to noise (primarily due to photon statistics) in the measurement of the wavefront, errors due to imperfect sensing and correcting of the wavefront caused by finite subaperture size, and a bandwidth error term due to the finite timestep over which the atmospheric wavefront is sensed. In the non-ideal case, there will be additional error terms, particularly for imperfect calibration of the system and for uncorrectable errors of the telescope optics. For a detailed discussion of these error terms, see Angel (1994), and Macintosh (2002).

For plausible EAO systems on a CELT-like telescope, the halo intensity is 10^{-6} to 10^{-7} of the peak intensity -- extrasolar planets will still be 10^2 times dimmer than the halo. This is not insurmountable; IR astronomers routinely detect faint galaxies 10^3 dimmer than the uniform sky background, for example. CELT's EAO ability to detect planets will therefore depend on the smoothness of the halo. On short timescales the halo is completely non-smooth and is composed completely of residual speckles. In a long exposure, the speckles move around and the image smoothes out. Speckles resulting from different error terms smooth out over different characteristic decorrelation times, with rapidly decorrelating speckles contributing much less to the final image noise than slow speckles. Noise due to wavefront sensing errors is completely random every update of the AO system (~1 ms) and hence has much less effect on final signal to noise than other error terms. Noise due to wavefront fitting error decorrelates more slowly though the exact timescale is uncertain. (See Macintosh, et.al., 2001. We have chosen to use the most pessimistic predictions.) Noise due to static effects such as calibration never decorrelates (though it may be reduced through PSF measurements.)

The table below shows performance for our baseline EAO system, described further in Section 9.5.1. The assumed atmospheric parameters match those in Section 9.2.2 with the exception of a better than

average r_0(0.5 μm) = 26 cm. Residual errors are expressed both in nm and in terms of the resulting noise (deviations from smoothness in the stellar halo) in a 2-hour exposure as a fraction of the peak intensity. CELT is capable of detecting a Jupiter-like planet orbiting a sun-like star 8 pc away, making a significant population of stars accessible for an extrasolar planet search, including several known radial-velocity planets.

For the baseline EAO system observing a m_v = 4.5 star at a distance of 8 pc with a companion at 5 AU, a Jupiter-sized companion, 10^{-9} dimmer than its parent star, would be detected at the 10-sigma level in a 2-hour exposure as shown in the following table.

Table 9-7. EAO error budget

Error terms	**Sigma (nm)**	**Residual halo noise in 2-hour exposure**
Measurement	31	8.7×10^{-12}
Atmospheric Fitting	11	5.8×10^{-11}
Bandwidth	7	1.3×10^{-11}
Scintillation	9	3.3×10^{-13}
Calibration	*5*	4.2×10^{-11}
Telescope	2	6.6×10^{-12}
Static Fitting Error	*2*	
Edge Effects	Section 9.5.1	
Total AO error terms	35	1.3×10^{-10}
Photon Poisson noise		5.0×10^{-11}
Total noise		1.4×10^{-10}

Telescope segment edge and diffraction effects

A crucial issue for EAO is the effect of residual phase discontinuities at segment edges. Both errors in segment shape and errors in segment tip, tilt, and piston result in segment-edge phase discontinuities. A continuous deformable mirror will be incapable of exactly fitting these discontinuities. As a result, even after correction, severe phase errors occur in a region of width d at the edge of each segment. These phase errors scatter light into a highly correlated pattern, resulting in a bright, dense, stationary speckle pattern (Macintosh, et al., 2001). Preliminary simulations indicate that for our baseline EAO system, CELT would require residual segment phase discontinuities on the order of 25-50 nm RMS wavefront error (compared to the 76 nm level Keck currently achieves [Troy, et al., 1998]) for these effects not to dominate. This requires significantly better individual segment shapes than Keck.

If the CELT primary does not achieve this level, we must suppress segment edge effects in other ways. One approach is a modified Lyot coronagraph, in which a combination of a focal-plane stop and a pupil-plane mask tracing the segment edges blocks diffracted light. The relationship between focal-plane stop size and pupil stop size in such a coronagraph (Sivaramakrishnan, et al., 2001) would require a focal stop of radius $\sim\lambda/d_{seg}$. If CELT segments are 1m in diameter , and we use λ = 1 μm, an occulting spot of radius 0.2-0.4 arcsec will be required, limiting the ability to see planets in close orbits. More advanced phase coronagraph concepts (Badouz, et al., 1999; Rouan, et al., 2000) may improve this and will be studied.

EAO illustrative design

Angel (1994) and Stahl and Sandler (1995) show representative EAO layouts. The EAO system would consist of a tip/tilt correction mirror followed by a conventional low-order deformable mirror and Shack-Hartman WFS. This will allow an initial AO correction, which will deliver the diffraction-limited image needed before the interferometric wavefront sensor can close its loop. It will also reduce the stroke requirements for the high-order DM. The visible light will continue on to a high order deformable mirror (see below) which is followed by a beam splitter where 80% of the light will be directed to a wavefront sensor. The remaining visible light will be re-imaged by a science camera (see Chapter 10).

EAO illustrative design

Number of deformable mirrors	2 (1 conventional + 1 MEMS)
DM conjugate height	0 m
Number of actuators	~ 3000 and ~ 200,000
Wavefront sensors	1 conventional + 1 interferometric
Controller –3db bandwidth	~ 1kHz

Wavefront sensing requirements

Initial wavefront control begins with the use of a conventional wavefront sensor to drive the lower order DM and achieve a diffraction-limited core. Once that is achieved, a Mach-Zehnder phase-shifting interferometer (Angel 1994) will be used to drive both DMs. This type of sensor measures the phase and amplitude of the wavefront directly. This reduces the computing requirements (see below) and minimizes measurement error. It will only operate efficiently on a near-diffraction-limited image, so the conventional sensor is needed to bootstrap the system. Such a WFS has been demonstrated open-loop on the sky (Colucci 1994), but further investigation and development is needed.

Deformable mirror requirements

The illustrative system requires two DMs: one for large-stroke low-order correction and another for the fine wavefront connection. The first DM will be used to insure the higher order DM does not saturate. This DM would require an actuator spacing of ~0.5 m at the primary mirror or ~3,000 actuators, stroke sufficient to correct the atmosphere, and update rates of ~1 KHz. This is only a factor of 2-3 times more than the largest DMs in use today and is reachable with state-of-the-art technology (M. Ealey, private communication).

The highest order DM requires an actuator spacing of 0.03 m at the primary mirror or ~200,000 actuators, and needs to support update rates of ~10 kHz, but with small stroke. Two possible technological solutions are spatial light modulators (SLM) and MEMS devices (Olivier 2001). Since these devices are highly desirable for other AO applications (e.g., horizontal-path communications and imaging) their development will be supported outside the astronomy community. However, astronomers must remain involved in this development to ensure devices can meet astronomical goals.

Computing requirements

The use of a direct phase measurement in the EAO system should considerably reduce the computing requirements over a wavefront slope sensor. The outstanding necessity is to create DM actuator commands given a phase map -- this may not be a one-to-one mapping since DM actuators typically have a spatial influence function that may overlap the influence of adjacent actuators. CELT EAO would have 200,000 actuators and a 10 kHz frame rate. Assuming an inverse-influence width of 5

actuators, this is 5 x 10^7 calculations/ms, or about 1500 times the calculation rate of the current Keck AO system. However, it is only ~20 times the calculation rate of the USAF visible-light AO systems currently in use. The algorithm can be highly paralleled.

Amplitude correction requirements

A first-order analysis shows that scintillation correction (at a good site) will not be necessary. Scintillation effects are less severe for our representative atmospheric profile than for that used in previous work (Stahl and Sandler 1995) and also are less proportionally severe for an extremely large telescope. This will be explored via simulation in the future. If scintillation correction is necessary, it could be carried out using either spatial light modulators (Love and Gourlay 1996) or using multiple deformable mirrors at different conjugate heights (Barchers 2001). The former approach is computationally much simpler.

Impact on other CELT subsystems

Telescope

Small phase errors introduced by dimples in the centers of Keck segments limit the high-contrast performance of Keck's current (non-extreme) AO system, and should serve as a cautionary example for AO/telescope high-contrast interactions. As discussed below, EAO performance will be limited either by segment edge phase discontinuities or by the coronagraphic mask needed to block these phase errors. The tradeoff is unclear between sensitivity loss due to the more aggressive Lyot blocking required for smaller segments, and sensitivity gain resulting from the improved surface figure of smaller segments. A detailed study of segment/coronagraph/EAO interactions should be carried out in the next design phase to ensure that the primary mirror does not limit high-contrast performance.

Instrumentation

A discussion of the instrumentation needs for EAO is presented in Section 10.2.4.

Site selection

EAO performance is a steep function of r_0; the performance predictions here assume a somewhat better than average Mauna Kea r_0 of 0.26 m at 500 nm. Reducing this by a factor of two increases the required AO system bandwidth by nearly a factor of two and requires target stars ~30-40% closer for a given S/N on a Jupiter analogue. The dependence on wind velocity and C_n^2 is much weaker; of all site parameters r_0 is by far the most important.

Outstanding EAO issues and future work

The design and performance of the telescope can severely impact the performance of a high-contrast imaging system. We need to understand these impacts so that informed decisions can be made regarding telescope design and error budgets. As one example, both static and dynamic wavefront errors at segment edges have a large impact on performance. Coronagraphic techniques can reduce some of these effects, but with a penalty in transmission.

Macintosh (2002) gives more detail of the issues discussed in this section. Key questions remain regarding the first order performance analysis used here. We need to better understand the time scales over which errors are averaged, the effects of systematic and static errors, and the validity of assumptions about the spatial characteristics of noise. Simulations will be needed to confirm the analytical expressions regarding these issues. It is also likely that experiments using existing AO systems and instruments will be needed to verify and develop the analytical calculations.

9.5.2 Ground-Layer Adaptive Optics (GLAO)

Introduction

For science requiring a wide FOV, it is possible to partially correct the atmospheric phase errors over large angles, while sacrificing the degree of correction obtainable over narrow FOVs. The concept, depicted in Figure 9-9, is to correct only for atmospheric turbulence originating near the ground, where its affect on incident wavefronts is common to many different directions. (In the extreme case, correcting only for turbulence arising exactly at the primary mirror would provide partial correction over the entire FOV of the telescope.) This technique is known as ground-layer adaptive optics (GLAO).

The scientific benefit of partial compensation is an increase in the sensitivity realized by the decrease in the image width and corresponding increase in peak irradiance level for each object in the (large) FOV. An example of the partial level of correction and the large useful FOV is shown in Figure 9-9.

Tomography error

The fundamental challenge of GLAO is the determination of the wavefront errors arising in the ground layer, and thus common to wide FOVs, in the presence of confusing turbulence layers at high altitude. This is a special case of the general problem of phase tomography, the reconstruction of the three-dimensional atmospheric phase aberration function from a set of two dimensional slices, with each guide star direction providing one slice.

Recent analytical covariance studies of the optimal configuration for ground-layer estimation have explored the dependency of tomographic error upon the maximum extent of guide star asterisms. For recent covariance calculations considering the problem of ground-layer wavefront sensing on a 5-m telescope (B. Mathews, private communication), the tomographic error is found to decrease with increasing guide star asterism spacing, reaching a minimum at approximately 17 arcmin diameter FOV; and a minimum of ~80 nm rms when using a Cerro Pachon turbulence model modified for increased resolution of the ground layer. For an asterism spacing beyond this, the geometry is such that the shear of the telescope pupil for the furthest separated beams, projected to 1km, begins to exceed the pupil diameter. It is expected that these geometry-driven results are scalable to CELT, although the dependence of tomography error upon telescope diameter has not yet been explored.

GLAO error budget

We adopt for the GLAO system's performance goal that the point spread function of a long-exposure image is reduced from the nominal 0.5 arcsec FWHM to 0.35 arcsec FWHM. To accomplish this, the GLAO system must reduce the integrated C_n^2 (h) value by 45% without degrading the image significantly by making excessive tomography error. We choose an error budget representing correction of a ground-layer containing 45% of the distributed C_n^2(h) to the level of 70% Strehl ratio at J-band (1.2 μm). Under these conditions, the encircled energy within 0.35 arcsec FWHM will closely approximate that produced by an atmosphere having 45% less integrated C_n^2 (h).

GLAO error term	**nm, rms**
Wavefront fitting	70
Tomography	80
Measurement	20
Controller bandwidth	41
Total error	119

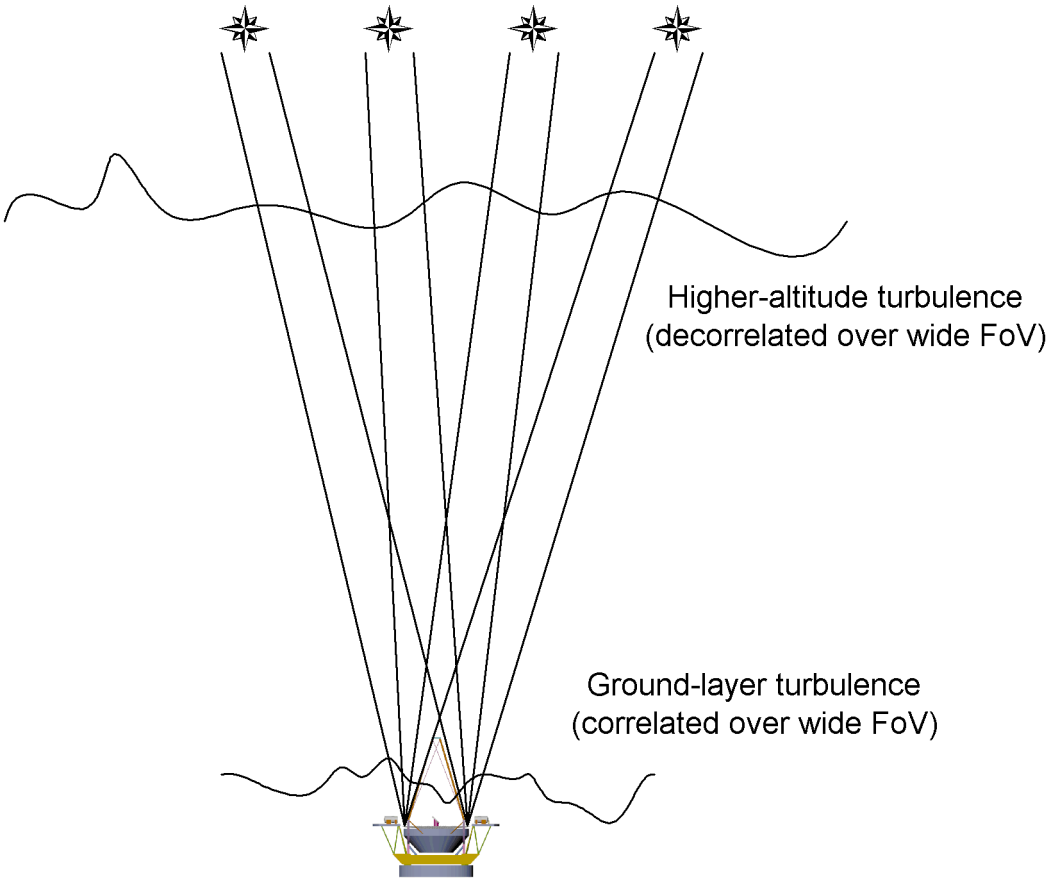


Figure 9-9. The ground-layer adaptive optics (GLAO) concept. Correction of errors arising from turbulence layers near the telescope provides a partial correction with a wide FOV. The technical challenge is disentangling the ground-layer from the higher altitude layers with sufficient fidelity.

GLAO illustrative design

A system having the following specifications will meet the budgeted residual wavefront error.

Number of deformable mirrors	1
DM conjugate height	200 m
Number of actuators	6,279
Controller –3db bandwidth	120 Hz
Number of guide stars	5
Diameter of guide star asterism	20 arcmin
Brightness of guide stars	180 photodetections / subaperture / exposure

NGS vs. Rayleigh beacons

Because of the very large FOV of the illustrative system, there are many sufficiently bright natural guide stars that can be exploited. Within a 20 arcmin square FOV, assuming a galactic latitude b = 30 degrees and longitude l = 90 degrees, there are expected to be 7 stars of infrared magnitude $m_K = 8$ or brighter (Wainscoat 1992), making an all natural guide star GLAO system feasible. However, this approach would require sharing infrared science light with a wavefront sensor.

An alternative approach is to project an asterism of low-altitude Rayleigh laser beacons as GLAO guide stars. Because we are specifically interested in sensing and correcting only the ground-layer, the usual drawback of the finite height of Rayleigh beacons instead becomes an advantage. In addition, the usual restriction of laser guide stars (that they do not measure tilt terms) is not important for GLAO, since global tilt errors for CELT will be much smaller than the 0.35 arcsec FWHM of the seeing-augmented point spread function.

Outstanding GLAO issues and future work

The concept of GLAO is new to the field of adaptive optics, which may be appreciated by the absence (at present) of any literature reference specifically addressing this technique. Researchers at Caltech and elsewhere are vigorously pursuing the analytical development of this mode of partially compensated imaging.

As described here, the GLAO system could be useful for wavelengths longer than 1 μm. To be effective at shorter wavelengths it will require a more complex system. Performance will depend critically on the C_N^2 (h) distribution.

The limitations of tomography error on GLAO remain the key issue for future investigation. The sensitivity of the technique to the vertical C_n^2(h) profile also requires further investigation. It is assumed for now that acoustic sounding or other techniques can provide real-time C_n^2(h) measurements of the ground-layer up to 1 km altitude (S. Flatté, private communication). Finally, better understanding of the theoretical and practical tradeoffs between NGS and Rayleigh LGS wavefront sensing of the ground layer are necessary.

9.6 Summary

To realize its full scientific power, CELT must extend the astronomically nascent field of adaptive optics into new architectures suitable for a 30-meter diameter telescope. The AO requirements exceed the state of the art implemented on 5-10-meter diameter telescopes in many ways. The total number of control channels is dramatically increased, and with it the high-speed computation requirements. The calibration and stability requirements are increased to accommodate the superb achievable resolution. Finally, the practical implementation of laser guide stars, barely explored at existing observatories, becomes significantly more challenging due to CELT's large aperture.

Despite these challenges, we believe not only that adaptive optics is feasible on CELT, but that the opportunity exists for significant scientific improvement over the first-generation AO systems deployed at major observatories during the 1990's. Better photometric accuracy, increased spectroscopic sensitivity, and greater imaging contrast are expected results of overcoming the new fundamental challenges of needing multiple laser guide stars and, most likely, multiple deformable mirrors. Enabling adaptive optics on CELT in order to realize this potential will require significant and sustained technical investment during the observatory design and construction phases.

During the CELT conceptual design phase, significant attention has been placed on identifying cost-saving "innovation factors." From one point of view, the use of adaptive optics is itself a source of innovation. AO-fed instruments, for example, can be made both smaller and simpler than seeing-limited instruments. *For unresolved, background-limited targets, CELT's sensitivity with full AO correction is equivalent to that of a 900-meter diameter primary mirror without any adaptive optics*.

Although not all science observations enjoy this gain, AO clearly provides a major innovation factor over historic telescope scaling laws, resulting in an excellent scientific return on investment.

For CELT, we must develop and demonstrate new technologies at a rate much faster than have past AO instrument programs. This will require concurrent engineering and development techniques. We must work with commercial partners to develop as-yet nonexistent component technologies while supporting their integration in testbed systems that will mitigate overall observatory cost and schedule risks. Finally, we must leverage ongoing national and international development of adaptive optics, where possible, to narrow the focus of CELT investment and defray development costs.

REFERENCES

Andrews, L., and R. Phillips. 1998. *Laser Beam Propagation Through Random Media*. SPIE: The International Society for Optical Engineering, Washington.

Angel, J.R.P. 1994. "Ground-based imaging of extrasolar planets using adaptive optics." *Nature* **386,** 203.

Babcock, H.W. 1953. "The possibility of compensating astronomical seeing." *Publications of the Astronomical Society of the Pacific* **65**, 229-236.

Bahcall and Soneira. 1980. "The Universe at faint magnitudes, I. Models for the Galaxy and the predicted star counts." *Astrophysical Journal Supplement* **44**, 73-110.

Barchers, J.D., and B.L. Ellerbroek. 2001. "Improved Compensation of Turbulence-induced Amplitude and Phase Distortions by Means of Multiple Near-field Phase Adjustments." *Journal of the Optical Society of America A* **18,** 399.

Baudoz, P., Y. Rabbia, J. Gay. 1999. "Achromatic interfero coronagraphy I. Theoretical capabilities for ground-based observations." *Astronomy and Astrophysics Supplement* **141**, 319.

Bauman, B., and R. Dekany. 2002. "Optical design issues for 30 m class multi-conjugate adaptive optics." CELT Report No. 24.

Colluci, D. 1994. Ph.D. Dissertation, University of Arizona.

Dekany, R. 2000. "A 1600 actuator upgrade for Palomar Adaptive Optics." SPIE Proceedings **4007**.

Dekany, R. 2002. "Integration time comparison for several mid-IR configuration options." CELT Technical Note No. 22 in preparation.

Ellerbroek, B., and F. Rigaut. 2000. "Scaling multi-conjugate adaptive optics performance estimates to extremely large telescopes." SPIE Proceedings **4007**, 1088-1099.

Fried, D. 1982. "Anisoplanatism in adaptive optics." *Journal of the Optical Society of America A* **72**, 1, 52-60.

Fried, D. 1995. “Focus anisoplanatism in the limit of infinitely many artificial-guide-star reference spots.” *Journal of the Optical Society of America A* **12**, 939-949.

Gavel, D. 2001. “Technology challenges to adaptive optics on extremely large telescopes.” in Proc. “Beyond Conventional Adaptive Optics” Conference, Venice.

Gemini Observatory. 2001. “MCAO for Gemini South.” Preliminary Design Review Documents, Section 5.2: “Laser System.”

Hardy, J.W. 1998. *Adaptive Optics for Astronomical Telescopes.* New York: Oxford University Press.

Hill, J.M., and P. Salinari. 2000, “The Large Binocular Telescope Project.” SPIE Proceedings **4407**, 36-46.

Lloyd-Hart, M., et al. 1998. “Infrared adaptive optics system for the 6.5-m MMT: system status and prototype results.” SPIE Proceedings **3353**, 82-93.

Love, G.D., and J. Gourlay. 1996. “Intensity-only Modulation for Atmospheric Scintillation Correction by Liquid-crystal Spatial Light Modulators.” *Optics Letters* **21**, 496.

Macintosh, B. 2002. “High-contrast adaptive optics for CELT.” CELT Technical Note No. 23.

Macintosh, B., S. Olivier, B. Bauman, J. Brase, E. Carr, C. Carrano, D. Gavel, C. Max, J. Patience. 2001. “Practical high-order adaptive optics systems for extrasolar planet searches.” to appear in SPIE Proceedings **4494**.

Nelson, J. 2001. “How Well Can the CELT Primary Approximate a Parabola?” CELT Technical Note No. 7.

Olivier, S.S. 2001. “Advanced Adaptive Optics Technology Development.” to appear in SPIE Proceedings **4494**.

Olivier, S., C.E. Max, J.M. Brase, D.T. Gavel, B. Macintosh, C. Carrano. 1998. “Direct Imaging of Extrasolar Planets.” *Proceedings International Workshop on Brown Dwarfs and Extra-Solar Planets*, ASP Conference Series **134**, 262. Rebolo, Martin, and Osorio, Eds.

Ragazzoni, R. 2000. “Adaptive optics corrections available for the whole sky.” *Nature* **403**, 54.

Ribak, E. 1998. “Alternative guide stars for adaptive optics.” SPIE Proceedings **3353**, 320-329.

Riccardi, A., N. Bindi, R. Ragazzoni, S. Espositio, P. Stefanini. 1998. “Laboratory characterization of a ‘Foucault-like’ wavefront sensor for Adaptive Optics.” SPIE Proceedings **3353**, 941-951.

Robin and Creze. 1986. “Stellar population in the milky way: a synthetic model.” *Astronomy and Astrophysics* **157**, 71-90.

Roddier, F., J. Gilli, G. Lund. 1982. “On the origin of speckle boiling and its effects in stellar speckle interferometry.” *Journal of Optics (Paris)* **13**, 263.

Roddier, F., L. Cowie, J. Graves, A. Songaila, D. McKenna. 1990. “Seeing at Mauna Kea – A joint UH-UN-NOAO-CFHT study.” SPIE Proceedings **1236**, 485-491.

Roddier, F., M. Northcott, and J.E. Graves. 1991. *Publication of the Astronomical Society of the Pacific* **103**,131.

Roddier, F., ed. 1999. *Adaptive Optics in Astronomy.* Cambridge: Cambridge University Press.

Rouan, D., P. Riaud, A. Boccaletti, Y. Clenet, A. Labeyrie. 2000. “The Four-Quadrant Phase-Mask Coronagraph. I. Principle.” *Publications of the Astronomical Society of the Pacific* **112**,1479-1486.

Sivaramakrishnan, A., C. Koresko, R. Makidon, T. Berkefeld, M. Kuchner. 2001. “Ground-based coronagraphy with high-order adaptive optics.” *The Astrophysical Journal* **552,** 397.

Stahl, S.M., and D.G. Sandler. 1995. “Optimization and Performance of Adaptive Optics for Imaging Extrasolar Planets.” *The Astrophysical Journal* **454**, L153.

Tatarski, V.I. 1961. *Wave Propagation in a Turbulent Medium*. New York: McGraw-Hill.

Taylor, K. 2002. “Large Aperture Fast Cameras.” CELT Technical Note No. 13.

Tokovinin, A., and M. Le Louarn. 2000. “Isoplanatism in a multi-conjugate adaptive optics system.” *Journal of the Optical Society of America A* **17**,1819-1827.

Tokovinin, A., and E. Viard. 2001. “Limiting precision of tomographic phase estimation.” *Journal of the Optical Society of America A* **18**, 873-881.

Troy, M., G. Chanan, E. Sirko, E. Leffert. 1998. “Residual Misalignments of the Keck Telescope Primary Mirror Segments: Classification of Modes and Implications for Adaptive Optics.” SPIE Proceedings **3352**, 307-317.

Tyler, G.A. “Merging: a new method for tomography through random media.” *Journal of the Optical Society of America A* **1**, 409-424.

Valley, G.C. 1980. “Isoplanatic degradation of tilt correction and short-term imaging systems.” *Applied Optics* **19**, 574-577.

Wainscoat, R.J., et al. 1992. “A model of the 8-25 micron point source infrared sky.” *Astrophysical Journal Supplement* **83**, 111-146.

Wizinowich, P., et al. 1996. “Adaptive Optics for Keck Observatory.” Keck Observatory Report No. 208.

Wizinowich, P., et al. 2001. “Lessons learned form Adaptive Optics on Mauna Kea.” The Mauna Kea Kamaaina Adaptive Optics Group, in Proc. “Beyond Conventional Adaptive Optics” Conference, Venice.

Chapter 10. Scientific Instrumentation

10.1 Introduction

Scientists at California Institute of Technology and the University of California have played important roles in the development of world-class instrumentation for astronomy, and have delivered many successful instruments over the years to the Palomar, Lick and Keck Observatories. Among those achievements for the Keck 10-meter telescopes are the LRIS multi-object spectrograph (Oke, et al., 1994) and the NIRC2 diffraction-limited infrared camera from Caltech; the HIRES (high-resolution echelle spectrograph) (Vogt, et al., 1994) and ESI (echellette spectrograph and imager) (Bigelow and Nelson 1998) from UC Santa Cruz; and the NIRSPEC infrared spectrometer (McLean, et al., 1998) from UCLA. Caltech and UC scientists have also been in the forefront of detector developments, innovative optical designs, and software development. They have established close ties with industrial partners and other major observatories worldwide. CELT is thus being born into a fertile environment for the development of innovative and challenging instrumentation.

10.1.1 Background and General Considerations

In order to take full advantage of an increase in telescope aperture (A), instruments need to scale proportionately to maintain their information gathering capacity (or $A\Omega$). This simple geometric rule is broken only when a decrease in the subtended solid angle on the sky (Ω) is afforded through sky aperture matching to finer atmospheric seeing conditions, aided where possible by adaptive correction of the incoming wavefront. The development of the 8-meter telescopes coincided with the realization that to fully capitalize on investments in very large telescopes, choosing sites with better atmospheric quality coupled with the development of adaptive optics (AO) was essential. Consequently some mitigation in the otherwise linear increase in instrument size was obtained. For CELT there seem to be no significant improvements in atmospheric conditions beyond those characteristic of Mauna Kea or the good Chilean sites, and hence a return to a simple scaling law for instrument size is a reasonable, first order assumption.

In an environment where telescope aperture considerations dominate, it is easy to forget that gains made over the last few decades are dwarfed in comparison to the gains afforded by the development of detectors. In the wavelength region from 0.3-1 μm CCDs have broad (~90%) QE response with very low (~$2e^-$ rms) noise. The near-infrared (IR) devices are not quite as good, however they are asymptomatically approaching perfection at a rate that is likely to significantly impact observational efficiencies on CELT timescales. The mid-IR régime also holds promises for such developments in the near future. However, for a variety of good physical (and commercial) reasons detector pixel-size is unlikely to scale with telescope aperture. Pixels, currently at <20 μm (at least in the important UV-to-near-IR region) will only get smaller with time while, with the development of edge-butting, quasi-contiguous formats can in principle grow to any size. For Nyquist (~2 pixel) sampling, camera f-ratios scale inversely with telescope diameter to a point where seeing-limited (~0.5 arcsec) sky apertures require impossibly fast (f/0.4) f-ratios, while field angles become the limiting factor to detector format. As far as we are aware, there is no solution to this double bind. Detectors on seeing-limited instruments will inevitably over-sample the data (unless pixel binning is used), while formats will be constrained by camera field angles rather than by individual detector areas. This is a new era where detector development no longer drives instrument design; rather, detectors will need to be customized for the extreme camera optics designed to optimize information packing onto the detectors.

Such considerations beg the question of what type of information is to be so packed: spectral, spatial, temporal, or a combination of all three. Pure imaging will have a limited but important role in the near-

and mid-IR; however, there is little doubt that spectroscopy will dominate in all wavebands. To fill the available detector real estate is then generally a choice between pure cross-dispersed HIRES, multi-object spectroscopy (MOS) at a more limited spectral resolution, or some hybrid of the two. A recent addition to this more traditional arsenal is the development of 3-D spectroscopy. These include integral field units (IFUs), tunable narrow-band filters (e.g., Fabry-Perots [FPs]), and imaging Fourier transform spectrographs (IFTSs). Taking the science case as our lead, we conclude that classical MOS techniques are favored for optical, seeing-limited spectroscopy while 3-D techniques, exemplified in particular by IFUs, are preferred in the multi-conjugate adaptive optics-fed near-IR domain. There is clearly cross-talk between these two broad alternatives, e.g., ground-layer adaptive optics (GLAO) and near-IR MOS, and seeing-limited IFTS work. However, we limit ourselves herein to those techniques that most clearly address the CELT science case.

10.1.2 Meeting the Science Requirements

To meet the science goals of CELT a suite of instruments is required. Some of these instruments will always be used under seeing-limited conditions and will make full use of the proposed 20 arcmin field of view (FOV). These instruments must be heavily multiplexed in object space so that hundreds or even thousands of spectra can be obtained simultaneously. The remainder of the desired suite of instruments will require some form of AO to overcome atmospheric turbulence and provide images at the diffraction-limit of the telescope. At the longer wavelengths of the mid-IR (8-27 μm), diffraction-limited performance can be achieved with low-order adaptive optics (LOAO). In the near-IR regime however, multi-conjugate adaptive optics (MCAO) will be required, and the corrected FOV will be, at most, 12 arcsec. A summary of the key science drivers and required instruments is given below.

For extrasolar planet-hunting and studies of disks around stars:
High-res optical spectrograph, near-IR coronagraph, mid-IR high-res spectrometer

For planetary and Solar System studies:
AO imaging and near-IR integral field spectroscopy

For nearby galaxies and the fossil record:
Optical high-res spectrograph, and AO imaging in near-IR

For Galactic nuclei, the Galactic center and black holes:
AO deployable integral field spectroscopy, AO near-IR imaging

For extragalactic studies and galaxy formation:
Wide-field multi-object optical spectrograph R~5000; wide-field multi-object faint object spectrograph (R~300); AO-based deployable IFU spectroscopy

The key parameters of these instruments are spatial resolution (either seeing-limited or diffraction-limited), spectral resolution ($R = \lambda/\delta\lambda$), wavelength range, and object multiplex advantage. Non-AO instruments require fast camera optics, and all of the instruments require state-of-the-art optical and IR array detectors with the largest possible formats. Table 10.1 provides a summary of the proposed instruments and Figure 10.1 gives a graphical representation of the parameter space covered.

Instrument Summary:

- CELT faint object spectrograph (CFOS): A low-to-medium resolution seeing-limited optical multi-slit spectrograph
- Medium-to-high resolution (MTHR) spectrograph permitting both high dispersion cross-dispersed echelle formats and medium resolution multi-fiber spectroscopy
- Near-IR MCAO-fed imager (AO Cam) allowing near diffraction-limited imaging in the J, H, K bands
- Deployable integral-field unit spectrograph (d-IFU) for spatially resolved, OH-suppressed spectroscopy at the near-IR MCAO focus
- Mid-IR camera/spectrograph (Mid-IR Cam/Spec) for imaging and low-to-medium resolution spectroscopy at the SCAO focus
- Extreme contrast adaptive optics (EAO) coronograph (optical coronograph) for diffraction-limited optical coronograph at the extreme AO focus.

Table 10-1. Summary of the properties of the proposed suite of instruments. An asterisk indicates that the spatial sample is the slit width (typically 2-3 pixels), except for the IR instruments where the spatial sampling (pixel size) is matched to the diffraction limit.

Instrument Name: Type	Wavelength range (μm)	FOV	Spectral R^n ($\lambda/\delta\lambda$)	Spatial Sampling*	Object Multiplex
CFOS: Lo-R spectrograph	0.40 – 1.0	20′	< 2000	0.75″	~1000
MTHR: Med-R /fiber	0.35 – 1.0	20′	< 18,000	0.75″	~700
MTHR: Hi-R echelle	0.30 – 1.0	7-20″	<200,000	0.75″	~10
AO Cam	1.0 – 2.5	40″	~5-100	3-7 mas	—-
d-IFU Spectrograph	1.0 – 2.5	2′	~4000	10-100 mas	< 80
Mid-IR Camera/Spectrograph	8.0 – 27.0	~1′	10-1000	3-7 mas	—-
EAO-Coronagraph	0.6 – 1.0	30″	~1000	8 mas	—-

10.1.3 Source and System Modeling

From a simple Poisson-statistical analysis, point-source limiting AB mags approaching B~27.9 (i.e., 5.3 mags below dark sky) should be achievable at a signal-to-noise ratio (S/N) ~5 for seeing-limited low dispersion (R~500) spectroscopy on CELT in four hours. In the same integration time, AO-fed (0.1 arcsec resolution) near-IR spectroscopy should be capable of reaching H~22.5 (~3.1 mags below the OH-clear sky continuum brightness) at an R~1000. These are just two examples of impressively faint limiting magnitudes; but they are dependent on a multitude of simplifying assumptions, which in turn are reliant upon the characteristics of the sky, telescope, instrument and source structure. However, they clearly demonstrate the need for very accurate (<0.2%) sky-subtraction and indicate the danger of ignoring systematic errors in the observational process. For CELT instrumentation it will be necessary to model in great detail the full optical chain from source to detector as part of the preliminary instrument design process, in order to characterize, quantify, and (where appropriate) mitigate systematic effects in the observations. Effects that will need to be included in the end-to-end analysis include sky structure and variability (spectral, spatial and temporal); changes in the telescope/AO point spread function (PSF); scattered light; instrument flexure; and general data reduction assumptions and procedures. Critical to this process is the character of the source structure itself. This latter issue, while particularly

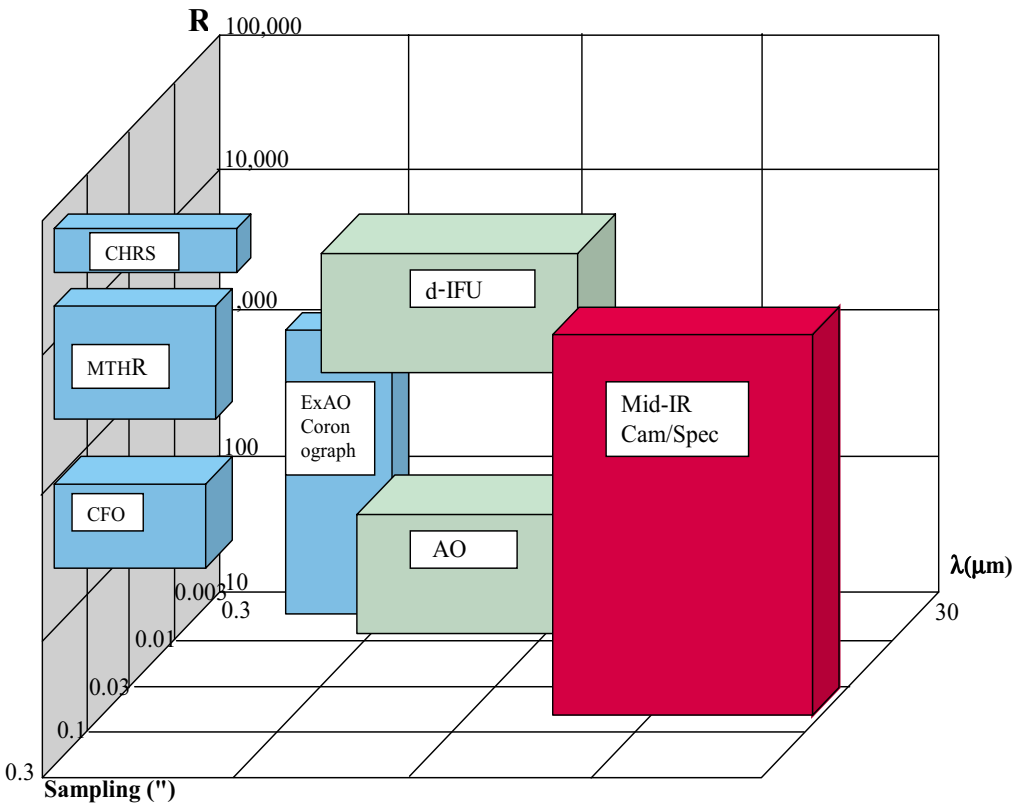


Figure 10-1. Illustrating instrument properties graphically in a three-dimensional plot of wavelength range, spectral resolution, and angular resolution (spatial sampling).

critical to the predictive process, is also the least quantifiable. Indeed, if we knew the structure of these faint sources, both spectrally and spatially, one might say that there would be no point observing them. Nevertheless, assumptions have to be made based on current observations, modeling, and analysis.

A general CELT instrumentation system modeling process has begun. We have constructed a set of generic source, sky, telescope, and instrument parameters that attempt to mimic the spectroscopic and imaging processes. The code is written in IDL and is run under an IDL GUI. It is designed to be freely expandable to embrace a wide variety of cases. Particular emphasis is given to predicting faint galaxy source structure as a function of redshift, including parameterized cosmological and evolutionary modeling. It is also designed to allow analysis of a wide variety of observational protocols with their systematic uncertainties.

The sources are typically modeled as 3-D (x, y, λ) cubes derived, i.e., from Hubble Space Telescope (HST) imaging. The output is generally in the form of a predicted spectrum with appropriately modeled shot noise contributions from the source, sky and detector.

Basic questions that can be addressed with this code include such issues as:

- optimal spectroscopic aperture sizes for classical MOS
- optimal integral field unit sampling (spatial and spectral)

- optimal spectral resolution for sky OH suppression
- flexure tolerance for sky-subtraction
- optimal pixel sampling
- sensitivity to detector noise characteristics

The parameter space is too large to permit general conclusions here; however, the goal is to supply a tool that can be used for any proposed instrument as part of its conceptual design, and beyond that as a broadly based S/N calculator for all CELT instrumentation. It is general enough to be useful for any ground-based OIR telescope. Baldry and Taylor (2001) give example analyses.

10.2 Initial Instrument Overview

Given the impracticality of adaptively correcting the incoming wavefront for $\lambda < 1$ μm over a 20 arcmin FOV, our baseline assumption is that only seeing-limited instrumentation will be employed in the optical domain, and that instrumentation at longer wavelengths will always be used with some form of AO.

10.2.1 Seeing-Limited Instruments

The demand for maximal throughput UV/optical spectroscopy, unencumbered by the inevitable losses of AO systems, is expected to remain strong, despite the fact that (in many cases) the necessary use of relatively large spectroscopic apertures will compromise sky-background-limited observations even for high dispersion (R~50,000) cross-dispersed echelle spectroscopy. In principle, an equivalent EAO-fed instrument would be far superior (smaller and cheaper). However, end-to-end (slit coupling) AO efficiencies of > 40% would have to be routinely achieved down into the UV in order to make such a proposal viable, and even then these gains would only apply for point sources.

Relatively large spectroscopic apertures (> 0.5 arcsec) are already necessary in order to integrate over marginally resolved objects such as distant galaxies (FHWM ~0.3 arcsec) (Taylor, et al., 1996). Integral field units, while offering the potential to reduce aperture size and hence sky background flux, are necessarily expensive in detector real estate. Consequently, in object (or wavelength) multiplex they are most effective on high surface brightness (line emission regions of complex sources that have, i.e., multiple point-like morphologies). Taking object multiplex into account, faint absorption-line galaxies, marginally resolved in good (< 0.5 arcsec) seeing, are still more efficiently studied using a classical MOS (multi-slit or multi-fiber) device (Baldry and Taylor 2001).

We are thus left with the need to investigate ways of using classical, cross-dispersed echelle spectroscopy and MOS in the seeing-limited optical régime. As noted in Section 10.1.1, any attempt to design spectrographs for seeing-limited work on CELT will require very fast cameras -- faster than f/1, if possible. Furthermore, the natural demand to maximize object and/or wavelength multiplex implies pushing camera designs to very large field angles (Taylor 2001a). In response to this need, Steve Vogt demonstrates in Section 10.5.1 a viable design for an 800 mm beam, ~f/1.3 camera operating from 0.3-1.0 μm. Given the inevitable cost and size of these cameras, our approach is to use them as a vast information collection resource that can be used for multiplexed observations in either the spatial or spectral domain. This leads to the MTHR concept for a single object, cross-dispersed high dispersion echelle combined with a multi-fiber intermediate dispersion spectrograph and associated fiber positioner. One clear science case requirement not easily satisfied with such a device is the important low dispersion, faint-limit optical spectroscopy normally addressed using wide-field, multi-slit devices such as LRIS

and DEIMOS (Keck) or VIMOS (VLT). Here the difficulties are not only in the scale of the optics implied by the 20 arcmin, f/15 focus but also in the requirement for atmospheric dispersion compensation (ADC) over the full field. The chosen approach to this challenge is to look at the possibility of an f/15-to-f/5 converter plus ADC feeding a VIMOS-like 4-shooter spectrograph. A preliminary design is presented.

10.2.2 Low-Order Adaptive Optic (LOAO) Instruments

Wavelengths longer than 2.5 μm are difficult to study from ground-based telescopes because of the thermal emission generated by telescope optics and Earth's atmosphere. Nevertheless, the exceptional angular resolution obtainable with CELT makes it desirable to observe protoplanetary disks that space missions cannot resolve. Since the diffraction-limit of CELT is ~70 milli-arcsec (mas) at 10 μm, it could be diffraction-limited at these wavelengths with only modest wavefront corrections. At these wavelengths, it is expected that the kind of AO systems available today (i.e., low-order systems) will be sufficient. One drawback will be the thermal background generated by the finite emissivity of the mirrors of the AO system, the telescope, and the secondary mirror support structure. Any mid-IR instrument will require either an optimal location (e.g., at a special focus preceded by an adaptive secondary) or a cryogenic AO system. In addition, the instrument will need to contain a cryogenic pupil stop that can be rotated and synchronized to the motion of the telescope pupil to eliminate unwanted thermal emission from the telescope.

The primary instrument envisaged for the SCAO mode is a mid-IR camera/spectrometer combination covering the wavelength range from 8-27 μm, but with performance optimized for the 8-13 μm window.

Key requirements of the design would be:

- a coronagraphic mode with occulting disk and rotating Lyot stop
- grism spectroscopy with R > 4000
- Nyquist sampling - broad wavelength range requires two cameras
- 34 mas/pixel at 10 μm; 69 mas/pixel at 20 μm
- final f-ratio: f/5.4 (demagnification = 2.8)
- FOV: 70.4 x 70.4 arcsec (154 x 154 mm at f/15) with $2k^2$ As:Si array

Detectors for this wavelength region are currently only 256^2 pixels and would be challenging to produce in the 2k format for applications other than spectroscopy because of the number of readout channels and very high speeds required.

10.2.3 Multi-Conjugate Adaptive Optic (MCAO) Instruments

This mode will be optimized primarily for the near-IR (1-2.5 μm). The MCAO system will be located on one of the Nasmyth platforms. With significantly more optical components than the LOAO system, emissivity and throughput will be compromised, and operation at $l > 2.2$ μm is not likely to be practical. The corrected FOV will probably be limited to ~2 arcmin. A direct camera for normal imaging at the diffraction limit of the telescope is desirable, although difficult to achieve in practice because of the huge number of pixels required. Since the diffraction limit is ~7 mas at 1 μm, it requires 2k pixels to Nyquist sample a field of only 7 arcsec. IR detectors of this size are currently being developed in a form that would potentially allow them to be assembled in a mosaic. For most applications, the camera should at least image the central ~30 arcsec of the MCAO field with Nyquist sampled pixels at 1 μm. For increased efficiency, the camera should observe more than one wavelength simultaneously.

For spatially resolved IFU spectroscopy, fundamental instrument parameters depend critically on the science goals. Given typical core sizes (~<0.3 arcsec) and surface brightness levels (<17 mag/arcsec^2 at H or K) of z >1 galaxies, spatially-resolved spectroscopy at the CELT diffraction limit will generally result in marginal S/N (Baldry and Taylor 2001). We assume, therefore, modest spatial samplings in the region 25-100 mas for this work. Furthermore, at K~22.5 the field galaxies surface density is ~20/arcmin^2 implying ~60 d-IFUs within the 2-arcmin diameter MCAO field. With subgalactic masses expected for the majority of such objects, spectral resolutions above R~3000 are required to gain kinematic information, R~4000 being perhaps a better value for optimal OH-suppression. On the other hand, a variety of science targets such as Solar System objects, protostellar disks, quasar hosts, the Galactic center, and nearby AGN also benefit from integral field spectroscopy, but do not require multi-object capabilities. These may be better suited to a "dense-pack," large-IFU configuration that can, in principle, use the same d-IFU spectrograph. The implementation is critically dependent on the chosen technology and will not be discussed further herein. Instead, two distinct technological approaches to spatially-resolved spectroscopy will be discussed.

10.2.4 Extreme AO (EAO) Instruments

Even in a "perfect" image of a star, stellar light is distributed in the diffraction pattern of the telescope and is not a concentrated point source. Coronagraphic techniques can be used to suppress telescope diffraction and render the image plane darker near the central star. In its simplest form, the Lyot coronagraph, a focal stop blocks a large fraction of the light by occulting the central few λ/D widths of the image. Light that is diffracted to a larger angle by passing close to a diffracting element (pupil edge, spider, or segment gap) passes the focal stop, but can be reimaged in a following pupil plane. In the pupil plane, this light is closely associated with the edges that caused the diffraction, and can be blocked with a Lyot (undersized pupil) stop. A trade must be made between the size of the occulting spot and the fraction of light blocked. For a 4 λ/D diameter stop, 50% of the pupil is blocked. At high Strehl, the apodisation of the focal stop with a smoothly varying transmission profile can significantly improve rejection of on-axis light. The Lyot coronagraph is robust and straightforward to implement, but suffers the disadvantages of limited rejection, low throughput, and complete obscuration of the central FOV. Nulling and interferometric coronagraph techniques can circumvent some of these limitations. Particularly promising is the phase quadrant coronagraph, which rejects all on-axis light from a perfect unobscured aperture. Its principal disadvantage is an extreme sensitivity to residual tip/tilt errors. The on-axis light appears outside the image of the telescope pupil. If the pupil mask is a mirror with a hole, the on-axis starlight can be refocused to a high bandwidth tip/tilt control loop to ensure sufficient centering on the phase mask. Such a coronagraph behind an EAO system would reject nearly all of the light in the diffraction pattern of the telescope. These novel coronagraphic techniques are presently being proven on existing high order AO systems.

After suppression of diffraction by a coronagraph, a science instrument focused on the science goals of the EAO system is required. For the primary science driver (direct imaging of extrasolar planets), different approaches are likely dependent on whether the position of the planet is unknown or known (e.g., via astrometric orbits from SIM or Keck interferometer). If the position of the planet is known, then it is not necessary to have a large number of spatial pixels. In this case, it would be possible to apply techniques to further suppress the light in the PSF and speckle pattern by taking advantage of properties of the planet (e.g., methane absorption, polarization); properties of the speckle pattern (e.g., wavelength dependent scaling, dark speckles); or both, with an integral field spectrograph, a simultaneous dual-imaging polarimeter, or a photon-counting dark speckle camera.

10.2.5 Ground-Layer AO (GLAO) Instruments

A new and potentially exciting possibility is that offered by ground-layer AO (GLAO). This is an AO technique that corrects only for the lowest turbulent atmospheric layer in order to deliver significantly improved seeing over a large ~5 arcmin FOV. As such it represents an important crossover technique that makes no attempt to deliver diffraction-limited imagery, but nevertheless corrects for much of the spatially correlated, low frequency power of the atmosphere (and telescope). For spectroscopy, where input aperture coupling and FOV are often at more of a premium than observing at the diffraction limit, GLAO promises very substantial gains.

GLAO is still very much in the conceptual stage, and so we can only make informed guesses as to its performance. For the sake of argument, we will assume that CELT will have a GLAO facility that will deliver a factor of ~2 or more improvement in seeing for $\lambda > 0.8$ μm over a ~4 arcmin diameter FOV. If this is the case, then a fairly simple adaptation of current near-IR spectrograph designs from Keck to CELT could be envisaged. As an example, we could migrate the proposed KIRMOS design to make it compatible with a CELT GLAO facility.

KIRMOS, as proposed, is a near-IR, wide-field, f/15-to-f/2.3 all-transmissive focal reducer for use as an MOS spectrograph/imager for Keck. It supplies a $4k^2$ HgCdTe array made from four edge-butted Hawaii-2RGs (Rockwell Science Center), giving it a huge 11 arcmin x 11 arcmin FOV. Its 175 mm beam permits R~4000 J, H, K (grism) spectroscopy when matched to a 0.75 arcsec slit. A very similar CELT GLAO instrument would require a relatively small perturbation to the KIRMOS optical prescription to give it a ~3.5 arcmin x 3.5 arcmin FOV. In this case a very similar optical system (CIRMOS) would produce R~4000 J, H, and K spectroscopy for 0.25 arcsec slits. In principle, object multiplexes of >>100 could be achieved with a traditional multi-longslit approach, while nod and shuffle (N&S) micro-holes could increase this by more than a factor of two.

10.3 Detectors

All astronomical instruments require the most sensitive photon detectors available. For wavelengths from 0.3-1.0 μm, the silicon CCD has traditionally provided the best choice since its invention in 1970. In the IR (1-30 μm), hybrid arrays employing silicon circuitry for the readout technology and various other materials for the light-sensitive part have revolutionized IR astronomy since their development in the mid-1980's. Nevertheless, the science requirements and diffraction-limited performance of an extremely large telescope will place stringent new demands on these technologies. Larger format devices with better noise performance will be required. Special detectors for adaptive optics need to be developed. In addition, the instrument designers should start to consider detector geometry as a design parameter, since focal plane arrays do not necessarily have to follow the layout of square pixels in a rectangular array. Along with the improvements in the detectors themselves, there also must be improvements in the electronic data acquisition and storage systems that operate them.

10.3.1 Optical

Silicon CCDs and CMOS devices: Scientific CCDs with formats of 4.5k x 2k and buttable on four sides are available today for assembly into large mosaics, the largest mosaic in development being the CFHT Megacam with 18k x 18k = 324M pixels with a mosaic of 36 devices. Furthermore, there is no obvious impediment to expanding mosaics up to more than a gigapixel. Improvements in the design of silicon CCDs (40-50 μm thick n-channel devices with MBE backside passivation, 300 μm thick p-

channel devices) are gradually leading to better quantum efficiency over the entire 0.3-1.1 μm range. However, the great change in index of refraction from 0.3 to 1.1 μm will always lead to a trade-off in selection of anti-reflective coating. Provided the rate of readout is relatively slow, noise performance levels in the 1-3e^- rms are fairly routine, but there continues to be improvement in amplifier design, with 4-5 e^- noise at 1 MHz possible today. Controllers and data acquisition systems must evolve, however, in order to handle the next generation of mosaics (see Section 10.3.5).

In the silicon imaging industry there is a move towards CMOS-based application-specific integrated circuits (ASICs) that employ on-chip signal processing. Charge coupling is dropped in favor of direct access, rather like hybrid IR arrays. The primary drawback of these devices for science applications is that they do not provide the 100% fill-factor of CCDs and the amplifier noise is higher. On the other hand, the appropriate use of ASIC technology can greatly reduce the external electronic equipment needed for operation and data capture.

CCDs are usually thought of as being rectangular arrays of square pixels. However, this geometry can be modified to optimize instrument design. Recently, a new kind of CCD was designed for use in curvature wavefront sensing for adaptive optics. With relatively straightforward changes to the pixel layout and charge movement, this device has been proven to outperform avalanche photodiodes at all magnitude levels, with much higher reliability and much lower cost. For the detection of laser guide stars used in CELT, the AO wavefront sensor detector can have a geometry that maximizes the S/N of the elongated laser spots. The development of detectors should proceed in parallel with instrument design, to utilize the flexibility in detector layout for optimal instrument performance.

CELT should also support the final steps in the development of one-electron noise amplifiers that operate at MHz pixel rates. The drawback of these amplifiers will be their small full well (20,000 e^-); however, this limitation can easily be overcome by placing a second non-destructive readout amplifier in series that has large full well. The two-amplifier implementation will enable the full dynamic range of CCDs (18-20 bits) to be accessed. Combining all of these features will provide one-electron readout noise with 18-20 bit dynamic range, at 1 MHz pixel rate. These nearly ideal performance levels can become available in a few years time if CELT provides the proper level of development funding.

10.3.2 Near-IR

HgCdTe Photodiode Hybrid devices: One of the best technologies to emerge for near-IR detectors is based on mercury-cadmium-telluride, or HgCdTe. HgCdTe has the unique property that the semiconductor band-gap can be "tuned" by adjusting the concentration of mercury. To date, the best uniformity and overall performance has been obtained with short-wavelength cut-off material (<2.5 μm). Detector arrays based on this material are effective from 0.8-2.5 μm when operated at liquid nitrogen temperatures (77 ^{0}K). Readout noise of less than 10e^- rms is possible and dark currents are <0.01e^-/s. More recently, the introduction of practical molecular beam epitaxy (MBE) is leading to the appearance of new devices with better overall uniformity and lower dark current. In addition, MBE can extend the wavelength coverage to ~5 μm for thermal IR work, or reduce the cut-off to 1.7 μm for non-thermal applications. The largest-format HgCdTe arrays are currently 2k^2 pixels. Rockwell Scientific, the primary supplier of these detectors, is investigating how to increase the format to 4k x 4k pixels, either by butting 2k x 2k arrays or fabricating a monolithic 4k^2 device (Rockwell Science Center). Some of this work is being funded by Next Generation Space Telescope (NGST), and devices for space applications are likely to contain ASICs to simplify the flight electronics. Thus, mosaics of 8k x 8k will likely exist within the current decade.

CELT should ensure that the technology matures by supporting the development of either four-side buttable 2k x 2k devices or monolithic 4k x 4k devices, as well as supporting a laboratory in California to test the prototype detectors.

InSb Photodiode Hybrid devices: Indium antimonide (InSb) has been an important detector material for IR astronomy because its band-gap makes it sensitive from about 0.9 μm out to a wavelength of 5.5 μm. Raytheon Infrared Operations has developed backside thinning and surface passivation techniques that result in photodiode arrays with excellent quantum efficiency. The current state-of-the-art in InSb detector technology is centered on the ALADDIN $1k^2$ array that has 27-μm pixels, although a new development (called ORION) is now under way for $2k^2$ detectors (Estrada, et al., 1998). With a standard anti-reflection coating, the average quantum efficiency across the 1-5 μm band is ~80% (with uniformity good to about 5%). In the best devices, dark current at 35 $^{\circ}$K is about 0.1 e^-/s/pixel and the readout noise is ~40e^- rms in a single correlated double sample. Using multiple non-destructive reads, it should be possible to achieve 10e^- rms. Like Rockwell, Raytheon is actively competing for a multi-element focal plane array of large dimensions for NGST. They are also developing a short-wave cut-off detector to compete with Rockwell's standard HgCdTe product.

Development of the InSb devices will need CELT support to ensure their availability when CELT requires them.

Near-IR devices have some undesirable features that will require significant research and development to reduce or eliminate. The primary concern is charge "persistence" due to traps in or near the pn junctions. This effect is seen in both InSb and HgCdTe devices.

10.3.3 Mid-IR

Impurity Band Conduction devices: Mid-IR arrays employ blocked impurity band (BIB) conduction. The generic term for these detectors is impurity band conduction (IBC) detectors. Silicon doped with Ga (cut-off ~17 μm), As (23 μm), or Sb (29 μm) can serve as the photon detector. The same or very similar readout integrated circuits can be bump-bonded to these arrays. In an IBC device, a heavily doped infrared-active layer is placed in contact with a pure (undoped) epitaxial layer (the blocking layer) and the overall thickness of the device is greatly reduced. The blocking layer is isolated by the use of an oxide layer from metal contact pads and the device is usually back-illuminated. Boeing (Anaheim) and Raytheon (Santa Barbara) are the main suppliers of these devices. Currently, the state-of-the-art for astronomy is an array of 256 x 256 pixels with a 40-μm pitch; devices with this format will be used on the Space Infrared Telescope Facility (SIRTF). There is no fundamental reason why the detector arrays cannot be applied to larger format readout devices (e.g., 1k x 1k), and both companies are proposing such formats for NGST and other space missions. On the time scale of CELT, there is good reason to expect a mature 1k x 1k technology, but 2k x 2k seems a stretch. Since the SIRTF and NGST detectors are being developed for space, they are low-background devices with low noise and low dark currents. These arrays are typically unsuitable for ground-based imaging unless the pixel size in arcsec on the sky is very small. These arrays may be very attractive for high-resolution mid-IR spectroscopy.

10.3.4 Others

There are other optical/IR technologies that may be of interest within the next 10 years. One interesting technology is that of superconducting tunnel junctions (STJ) which are essentially photon-counting

detectors with the ability to discriminate the energy of individual photons. An array of STJ devices would constitute a combined camera and spectrograph that do not involve any dispersion device (de Bruijne, et al., 2001).

The development of STJs requires a very large percentage of the funds that are being provided by the space agencies, and at present it is beyond the capability of CELT to significantly affect the progress of this technology. However, the CELT team should keep close watch over progress in this area.

10.3.5 Detector Controllers and Data Acquisition Systems

The next generation of array detectors will drive the development of new and more efficient controllers and data acquisition systems. The UC/Caltech/CARA communities are already engaged in an effort to coordinate these developments in order to gain economy of scale and more uniformity among observatories and institutions (see Web reference 1). To ensure that these critical systems exist on the timescale of the detectors and the instruments, it will be necessary to begin soon to research and develop them. This should be a major goal for the next phase of CELT.

The basic requirement for the next generation of electronics and data acquisition system is to achieve detector-limited performance for as large a mosaic as desired by the instrument. Some of the features of the next generation electronics are:

- array-limited performance
- system noise ~ 1e- rms at nominal readout rates
- scalable
- flexible (for different arrays)
- modular: upgrade path
- compact and rugged
- handle arrays with a very large number of pixels (gigapixel and more)
- handle a large number of outputs (minimum of 64)
- very high output readout rates possible (up to 4 MHz per channel)
- low-noise preamplifier per output channel
- programmable gain (1-20)
- fully programmable voltages and electronic gain
- 16-, 24- and 32-bit-wide data paths
- 16-bit digitization at 1-2 MHz
- large number of clock lines with high/low levels positive or negative (voltage range as required by the devices)
- large number of dc voltages: levels positive or negative as needed by the detectors
- programmable readout and digitization waveforms
- Fowler sampling for IR arrays
- sub-array readout mode
- variable readout rate

Development of these systems needs to start right away, or there could be a major bottleneck in the production of instruments.

10.4 Multi-Object Techniques

10.4.1 Multi-Slits

Any wide-field multi-slit device for CELT is likely to have a massive input focal plane in the order of ~1-m diameter or larger, and it is in this context that technologies have to be judged. Currently there are three broad approaches:

Pre-machined slit masks: This is by far the most common and is regarded here as the default option. Issues regarding materials, cutting procedures, and cryogenic deployment have been widely studied (Oke, et al., 1994) and successful solutions implemented (see Web reference 1). The off-line nature of the process, requiring preimaging and mask design followed by machining (at significant cost), is often regarded as a drawback. It does, however, ensure the use of disciplined observing procedures that can have efficiency benefits when compared to the on-line decision-making observing model.

Automated mechanical slits: Moveable slits have been suggested on a number of occasions (including for NGST) and were implemented most notably in the VLT's FORS-1 instrument (see Web reference 2) and the Palomar double-beam spectrograph. They have the advantage of permitting on-line observations but are geometrically constrained in both the number of slits that can be deployed and also in the details of their configuration. While they are much more compatible with remote and automated observing procedures, their limitations preclude them from serious investigation (as evidenced by the fact that they were not implemented on FORS-2).

Micro-mirrors and -shutters: These newer technologies are under investigation for Gemini (Moseley, et al., 2001) and in active development for the NGST (MacKenty, et al., 2000). However, the meter-sized scale of CELT MOS focal-planes and the inherent relative inefficiency (< 80%) of the technology make this otherwise attractive option non-viable at this stage.

10.4.2 Multi-Fibers

There are four generic types of fiber positioner that have been developed for astronomy:
1. Manual plug-plate systems (Barden, et al., 1993)
2. MX-type fisherman-around-a-pond systems (Hill, et al., 1986)
3. AutoFib robotic placement of magnetic buttons on a field-plate (Taylor, et al., 1997)
4. Echidna device with articulated spines to carry the fibers (Gillingham 2000)

All have their strengths and weaknesses; however, manual plug-plate is clearly ruled out for CELT on logistical grounds while Echidna, still under conceptual development, is matched only for fast f-ratio prime-focus use. MX-type and AutoFib are the only technologies appropriate for deployment on the Nasmyth platform of a large optical telescope. While MX-type has the advantage that all fiber probes can be articulated in parallel allowing rapid field configurations, it is only the AutoFib technology that permits relatively unconstrained fiber-to-target allocation. The most well-developed and ambitious fiber positioners to date, the AAO's 2dF system (see Web reference 4) and the VLT's OzPoz (see Web reference 5) (designed and built by the 2dF team), work around the problem of slow fiber configuration by employing double-buffering. Here the fibers are robotically positioned on one field-plate while the other is being used for observations. Once the first field has been observed and the second fiber field configured, the telescope is moved to acquire the second field while the first is then re-configured by the positioner for the next fiber configuration. In this way no time is lost.

10.4.3 Multi-Slicers

There are currently three broad technological approaches to integral-field image slicing: bare lenslets (see TIGER/OSIRIS, Web references 6 and 7); fibers (usually coupled with lenslets, see as an example the AAO's SPIRAL [Kenworthy, Parry and Taylor 2001] development, and below); and slicers (see 3-D/AIS [Web reference 8; Content 1997]). In principle, slicers offer the most efficient means of reformatting 2-D information into a spectrograph (Content 1997). The difficulty is to solve the problem of their deployment across a telescope focal surface. This issue has been most recently addressed by the GIRMOS design study, a collaboration between the UK's ATC and the AAO that studied the pros and cons of the fiber and slicer approaches in the context of a near-IR d-IFU instrument for Gemini. This joint study was able to convincingly demonstrate that viable instrument concepts could be defined for both technologies. It also showed that the choice must be based not only on theoretical throughput and reformatting efficiency estimates, but also (and more critically) on the practicality and cost of the proposed technology.

Fiber slicers (bare or lenslet) can employ standard positioner technologies 1, 2 and 3 above, or pick-off mirrors as described below. On the other hand, mirror slicers require some form of deployment concept. One of these, independently arrived at by the ATC and AAO groups, employs multiple pick-off arms (m-PoAs), each of which is designed to patrol a sector of the focal surface in such a way as to permit efficient PoA-to-target allocation. Each PoA contains reimaging optics, and is elbowed to allow for a constant optical path length between input and output foci independent of its position in the field. In this manner, a PoA can access any field position within its scope of motion while delivering a relayed image onto a fixed IFU stationed appropriately at the input, thence to a fixed location spectrograph. The whole system (m-PoA-to-slicer-to-spectrograph) has to rotate as a body on the Nasmyth platform, unless a large 250 mm diameter K-mirror image rotator can be envisaged as part of the MCAO system.

10.4.4 3-D Devices

Tunable filter Fabry-Perots (FPs): These have seen a minor renaissance in recent years following the earlier FP imaging spectrographs (Atherton and Taylor 1982), through the development of ultra-narrow gap etalons combined with CCD charge-shuffling synchronous with l-switching (Bland-Hawthorn and Jones 1997), to produce reliably fluxed, Poisson-noise limited, narrow-band imaging. Pioneering work at the AAO has found a broad spectrum of scientific usage for such techniques; however, the demand is likely to remain secondary to that of classical spectroscopy. Nevertheless, a relatively modest commitment of resources could generate (for example) a very valuable MCAO facility for near-IR, diffraction-limited, tunable filter ($\lambda/\delta\lambda$ ~10-50) observations of line emission regions (galactic or cosmological).

Imaging Fourier transform spectrographs (IFTSs): Almost the same could be said of IFTSs, except that their potential is more for the optical rather than the IR domain, where the OH contamination would be fatal. This statement contrasts with the early development of IFTSs for the IR, where detector noise levels are higher and control of path delay is less critical (see Web reference 8). Developments currently underway (see Web reference 9) are targeted at much finer and more stable control of the beam combination, with a view to exploring the viability of the technique in the quieter optical domain. If successful, the IFTS will be far superior to the FP for both narrow- and broadband tunable-filter imaging, given its much cleaner instrumental profile. However, there is some way to go before such instruments can be seriously considered for the current generation of large telescope.

10.5 Detailed Instrument Designs

10.5.1 Seeing-Limited Instrumentation

High-resolution spectroscopy (MTHR's cross-dispersed echelle mode)

Figure 10-2 shows a detailed plan view of the proposed high-resolution spectrograph. MTHR's high-resolution capability (Vogt 2001), like UVES, is supplied by a dual-beam instrument consisting of two parallel (and functionally identical) red and blue arms. Each arm is a grating-cross-dispersed echelle spectrometer in UVES dual-white-pupil configuration, with a scaled-up HIRES-style camera feeding a focal CCD array of $8k^2$, 15-μm pixels. The blue side is coating-optimized for 0.3-0.55 μm, while the red side is optimized for the 0.45–1.1 μm region. However, since the HIRES-style cameras and mirror-collimators are almost totally achromatic, both sides can cover much larger wavelength regions (a factor which will prove very useful for fiber-feeding, as discussed later). A dichroic mirror splits the beam into red/blue arms, and multiple dichroics can be provided to fine-tune the crossover wavelength.

The converging f/15 beam from CELT passes through an atmospheric dispersion compensator (ADC) and I_2 cell on its way to the nominal f/15 focus. For very high-resolution (R >100,000) work, an image-slicer would be placed at this focus if desired. Following the f/15 focus is a fused-silica TIR image derotator, with a CaF/fused-silica doublet cemented to its output face. The doublet collimates the beam and produces a pupil 0.2 m downstream. Just behind this pupil, a dichroic splits the beam, reflecting the blue and transmitting the red. The red beam is then reflected off a small silver mirror immediately behind the dichroic. Separate red-side/blue-side CaF/fused-silica doublets focus each beam, presenting an f/10 beam into each side's slit. Each slit will have a TV camera for image acquisition and guiding. Each slit has a usable length of between 7 and 20 arcsec, and a small number of fibers (< 10) or a multi-slit plate could be inserted at the slit if desired.

Following the light path through the red side only (both sides are identical), the diverging f/10 beam from the slit passes under a steeply-tilted face-down echelle grating, off a small fold mirror, and then into an off-axis (5°) parabolic mirror that collimates the beam and sends it into the echelle. This collimator also produces a white-pupil near the echelle, keeping the echelle as small as possible. The echelle is an R-4, used in quasi-Littrow mode with a ~2^o out of plane angle. This echelle is quite large by present standards, 1.0 x 3.5-m in size, and would be a 5 x 8 mosaic of the largest presently available echelles. The quasi-Littrow-diffracted beam from this echelle passes back to the collimator and then back towards the first fold-flat, narrowly missing the flat, and thence to a flat transfer mirror. At this transfer mirror, all the echelle orders are stacked on top of one another, and the imagery is not particularly good due to aberrations from the collimator. The transfer mirror relays the beam to a second off-axis parabolic collimator, which is an identical match to and co-linear with the first collimator. This second collimator largely corrects the aberrations from the first, and forms a second white-pupil in the system near the grating cross-disperser. The cross-disperser then diffracts the light into a HIRES-style catadioptric camera, which uses two large fused-silica corrector lenses, a large fast spherical mirror, and a singlet fused-silica "field flattener," with a CCD array and its LN2 dewar at the internal focus.

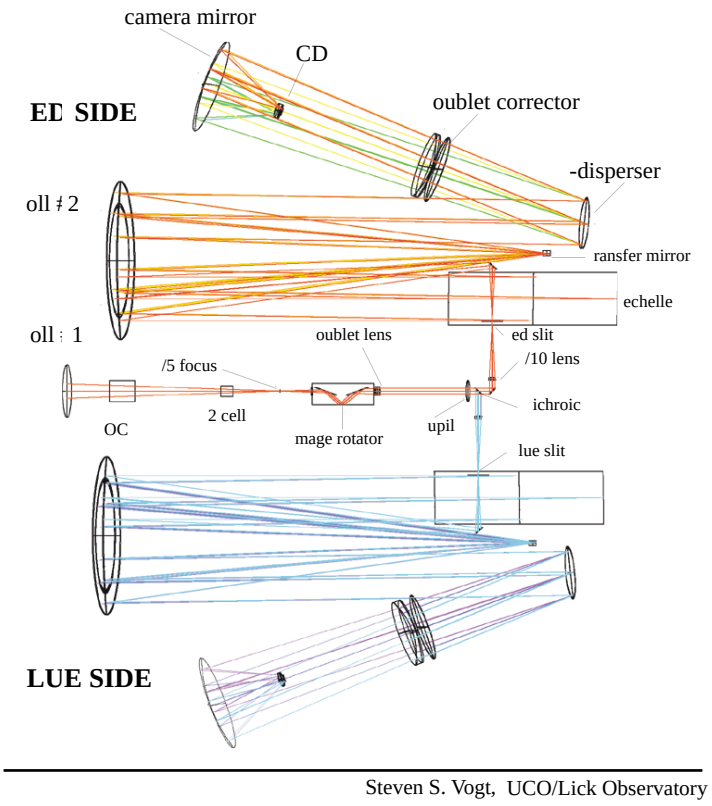


Figure 10-2. Detailed plan view of MTHR's high-resolution mode

Moderate resolution multi-object spectroscopy (MTHR's fiber-fed mode)
The large and expensive cameras, gratings, and two large CCD arrays in MTHR can also be used to good advantage for moderate resolution fiber-fed MOS, providing resolutions in the 2000-20,000 range for hundreds of objects at a time (Vogt 2001). Figure 10-3 shows one possible scheme. Here, what is essentially a HIRES-style camera (used in reverse) serves as a fiber-collimator, accepting light at f/4.5 from the outputs of a line of fibers. (These fibers are fed at f/5 from CELT, so allowance for focal ratio degradation is already built into this design.) The collimated beams pass through two large fused-silica corrector lenses, and then off the cross-disperser and into the normal MTHR camera. This fiber collimator sits alongside the MTHR camera, but on the other side of the camera axis from the echelle so as not to interfere with the light path of the high-resolution mode (and to eliminate any need for moving or stowing of optics). Both arms of MTHR could have such a fiber feed.

In Figure 10-3 the red beam is from a fiber at the bottom end of the line of fibers, while the blue is from the top end. These two beams land respectively at the upper and lower edges of the $8k^2$ CCD array. There is room here for a ~332 mm long fiber stack, or a maximum of about 2612 (touching) 87-μm core fibers (assuming 20-μm cladding).

The cross-disperser for this mode *could* be the same grating used for the high-resolution mode, however that grating would need to be rotated 180° about its normal to stay on the same side of its blaze. It is likely that other optimized cross-dispersers would be used, mounted on a cross-disperser carousel alongside the high resolution cross-disperser.

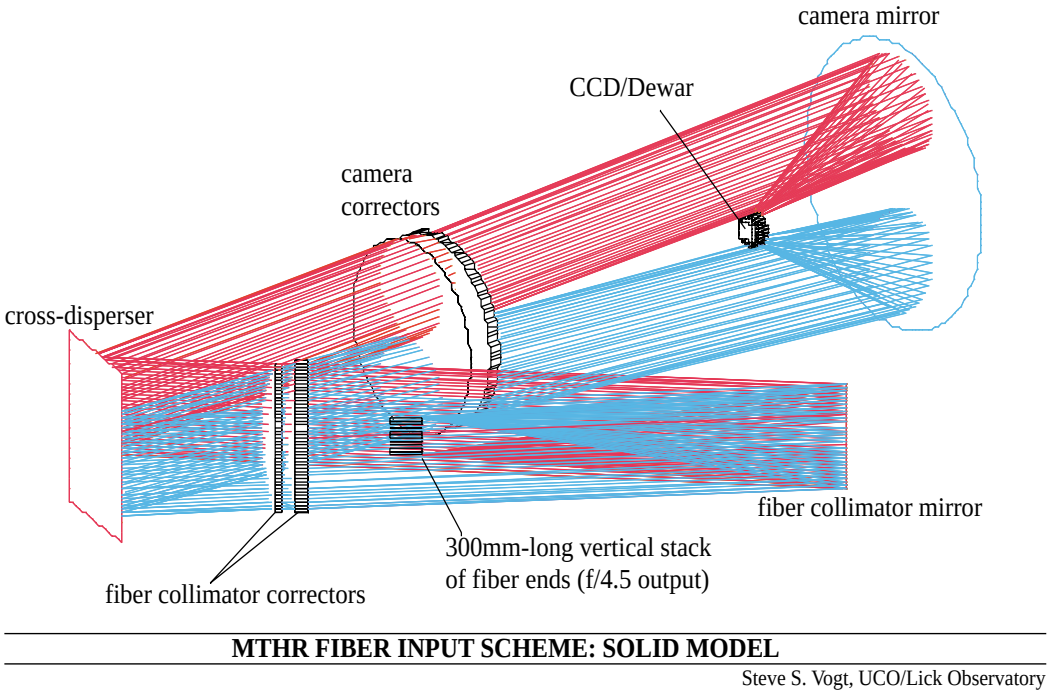


Figure 10-3. Moderate resolution multi-object fiber-input mode

The fiber-feed concept naturally makes for an *enormously* flexible design, and there are many possibilities involving different cross-dispersers, different configurations of fibers (e.g., multiple deployable single fibers; and SPIRAL 7-, 19-, and 37-hex IFUs), fiber feeds from a GLAO focus, etc. As a simple benchmark, the present design accommodates enough fibers for ~70 SPIRAL IFUs (see Web reference 11). Each 37-fiber IFU comprises a 0.73-arcsec-diameter bundle of 0.12 arcsec (87 μm) fibers deployable across CELT's 20 arcmin diameter FOV using an OzPoz or similar fiber positioner. A 300 gr/mm cross-disperser yields a resolution of ~4100 and wavelength coverage of 2800A. The 87-μm wide "fiber slit" of each IFU projects to ~38 μm at the CCD, close to the Nyquist optimum for 15-μm pixels. Higher spectral resolution can also be obtained (trading off spectral coverage for spectral resolution) by using finer-pitched gratings. Grating rulings of 600 and 1200 gr/mm would yield R~12,000 and 24,000 respectively. In this way, the full span of intermediate resolutions from ~2000 to 24,000 could be addressed in the multi-object mode.

Approximately 70 objects would be obtained in one shot with this setup, or ~35 if space on the CCD is left for an N&S technique (Glazebrook and Bland-Hawthorn 2001) for better background cancellation. Moreover, both arms of MTHR could have such a fiber feed, thereby effectively doubling either the wavelength coverage or the number of objects. It is also interesting to note that either fiber-fed arm could be used at almost any wavelength (within the limitations of their respective AR and reflective

coatings) since the HIRES-style cameras (and collimators) in both arms are almost completely achromatic and highly transmissive over a much broader wavelength region. Thus, for example, both arms could deliver spectra from 4000Å to 7000 Å, accommodating 70 SPIRAL IFUs each, for a total of 140 objects.

Adaptation of MTHR for fiber spectroscopy

Introduction

In this section we develop Steve Vogt's suggestion of feeding the MTHR cameras with fibers. As was discussed in Section 10.2.1, the requirement to cover a 20 arcmin FOV for intermediate dispersion spectroscopy on CELT is extremely difficult for traditional multi-slit systems. Thus fiber technology, as developed for wide-field astronomy over the last two decades, has to be seriously considered. Early experience with the efficiencies of such systems was discouraging. However, improvements in fibers, together with a much deeper understanding of how to optimize throughput and photometric stability for the newer systems (in particular, 2dF [Taylor and Gray 1994], FLAMES [Pasquini, et al., 2000], and SPIRAL [Haynes, et al., 2000]), have led to end-to-end fiber system losses of <20%, comparable to the standard slit-jaw losses in a multi-slit spectrograph. With sensitivity goals of <5 magnitude below sky, all spectroscopic instruments are going to require strict control of systematic errors. Traditionally, fiber instruments are characterized by ~2% sky-subtraction errors. The development of charge-shuffling (Glazebrook and Bland-Hawthorn 2001), in consort with aperture nodding for CCDs (the N&S technique) and analogous techniques for the near-IR domain, have the potential to establish almost perfect, Poisson-noise-limited sky-subtraction. While it is true that for N& S work < 50% of the detector real estate and only 50% of the exposure time is allocated to on-source integration, this is to be compared with the traditional long-slit MOS approach where factors of > 4 in detector real estate are dedicated to sky (Glazebrook and Bland-Hawthorn 2001). Taylor(2001) details the case for fibers on CELT.

Aperture matching to the spectral resolution requirements

With fiber systems there is little freedom to customize input aperture size to prevailing atmospheric conditions or to the intrinsic object size. It is therefore necessary to establish a suitable compromise. Point source observations clearly call for smaller apertures; however, previous studies (Taylor 1996) have focused on optimization for distant galaxy work. We now have the tools for refining such a choice (see Section 10.1.3). For argument's sake and as a suitable compromise, we will adopt an aperture of 0.75 arcsec, somewhat biased for marginally resolved galaxies. It is unlikely that a proper optimization, if one can be defined, will produce a significant perturbation from this figure unless choices in fiber diameter can be facilitated (see later in this section).

Using the established fiber-IFU technologies, exemplified by the AAO's SPIRAL project, we can invoke a hexagonally formatted fiber-slicer aperture (Kenworthy, Perry and Taylor 2001) as an input to MTHR. The finer the slicing, the higher the resolution until the ~2-pixel resolution limit is reached. (While it is possible to recover finer resolution through sub-pixel dithering we are presently ignoring these possibilities, since the systematics for such spectroscopy are not yet well understood.) For the MTHR capability as described in 10.5.1, a 37-hex fiber-slicer was adopted as an example. Here we analyze the more general trade-off in object multiplex ($M_{\#}$) with slicer format under more optimized assumptions (Taylor 2001c): specifically, for a camera f-ratio of ~f/1.1 (close to the limit of speed for such systems [Taylor 2001a]), core:clad ratios of 1:1, and a 1-pixel separation between each object spectrum.

If we examine the effect of different hexagonal slicing choices as shown in Table 10-2, we see that diminishing returns in R are encountered beyond the 7-hex arrangement due to the fibers approaching the two pixel limit. If the requirements were heavily biased towards object multiplex rather than spectral resolution, one could use a single unsliced fiber aperture. However, with the possibility of a low resolution multi-slit MOS capability (see Section 10.5.1), the choice of a 7-hex fiber slicer would seem to offer the best intermediate dispersion (R < 15,000) compromise, giving object multiplex gains of ~140 in each MTHR camera (assuming the N&S mode of observing).

Table 10-2. Examples of fiber-slicer options

SLICER (0.75 arcsec)	**FIBER DIAM.** (μm)	$\mathbf{R_{Octave}}$ (λ_0~600 nm)	$\mathbf{R_{1200}}$ ($\Delta\lambda$~110 nm)	$\mathbf{M_{\#}}$ (2-arms+N&S)
Single fiber	600	1400	5400	700
7-hex	225	3600	14,500	280
9-hex	140	5400	18,500	170
37-hex	100	5400	18,700	120

Fiber positioner (CfP)

We have assumed here that CELT's fiber positioner (CfP) will be based on that of OzPoz (see Web reference 5), the Nasmyth positioner for the VLT FLAMES facility (see Web reference 12). Apart from up-scaling the design, very little needs to be changed in the basic concept. Here we show four field-plates to allow for a variety of options, including differently-sized fibers and fiber-based d-IFUs. As shown in Figure 10-4, each field-plate could in principle support ~1800 fiber buttons, well in excess of the MTHR concept (even in its low dispersion, $M_{\#}$ ~700 mode). Hence a variety of options could be supported, including different fiber apertures, different slicing geometries, non-N&S ($2M_{\#}$) modes, not to mention a range of deployable fiber-based IFUs (d-IFUs).

The main difficulties with such a concept are associated with the massive ~2.6 m diameter, f/15 focal plane. On the smaller 8-10-m class telescopes, both non-telecentricity and atmospheric dispersion effects can be optically corrected over a wide field; this is simply not possible for CELT, where we have to devise alternative strategies (Taylor 2001d).

1. **Non-telecentric feed:** In order to maintain good focus over the 20 arcmin FOV, the CfP field-plate on which the fiber buttons are held has to be profiled to match the ~5.6 m radius of CELT's f/15 focal surface. However, with a telescope pupil ~60 m back, the off-axis principle ray's incident on each fiber button will not be orthogonal to the field-plate nor to the fiber button's axis of symmetry. At the edge of the field this angle is ~14°, quite unacceptable for fiber feeds. In order to ensure that the fiber-optical axis is parallel to the principle ray at all field positions, we suggest an articulated button, as shown in the inset to Figure 10-4. Here the button is moved down to the field-plate in a direction orthogonal to the R-q robot, whose R-axis has its center of curvature coincident with the telescope pupil (i.e., ~60 m). The button will then seat itself on the curved field-plate while the upper part of the button, holding the fiber, will stay parallel to the principle ray.

2. **Atmospheric Dispersion Correction (ADC):** Atmospheric dispersion correction is essential in maintaining good throughput for multi-fiber (or multi-slit) low dispersion spectroscopy (Taylor

and Gray 1990). In order to correct for this, each fiber button will require its own ADC. A conceptually simple solution would be to mount a Risley doublet prism pair on each button allowing zero-deviation correction into the fiber. In order for such miniature ADCs to not constrain nearest-neighbor button-to-button distances, they must have diameters ~5 mm and be individually oriented by the R-q robot head before placement. This represents a challenging and expensive, but not impossible, requirement.

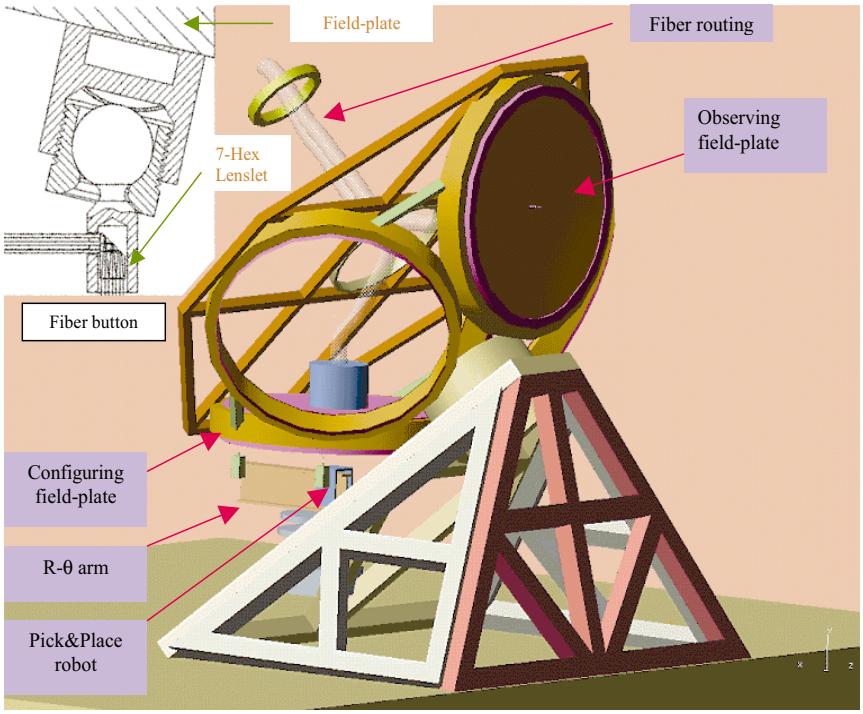


Figure 10-4. CfP, based on the OzPoz model (see Web reference 5), as mounted on the SL Nasmyth platform Shown inset is the concept for retaining a telecentric fiber feed while locating the fiber button on the spherical focal surface.

Low Dispersion Multi-Slit Spectrograph (CFOS)

f/15 to f/5 wide-field converter + ADC: We explore here the possibility of achieving low dispersion (R < 2000), wide-field (~20 arcmin), seeing-limited (0.75 arcsec slit-width), multi-slit spectroscopy on CELT through the use of a massive focal reducer, which acts as a feed for a four-shooter spectrograph/ imager. The original concept for the focal reducer (Jones and Taylor 2001) (the three-mirror focal reducer, or 3MfR) was produced by Damien Jones (Prime Optics, Qld., Australia). The preliminary design folds the beam vertically using a large atoroidal field mirror at the native 2.6-m diameter f/15 focus. The beam is then reimaged to a more manageable ~870 mm focus by a triple-mirror anastigmat (TMA) to produce an f/5 image surface whose distortion is stable to field rotation. There is sufficient space between the last element of the TMA and the input focal plane of the spectrograph to supply a classical Risley prism atmospheric dispersion corrector (ADC) whose diameter will be well under ~1 m (cf. 2dF's ADC, Web reference 12). One possible layout is shown in Figure 10-5. It will be noted that this arrangement has the spectrograph looking upward and rotating about a vertical axis, an ideal configuration for such a massive instrument.

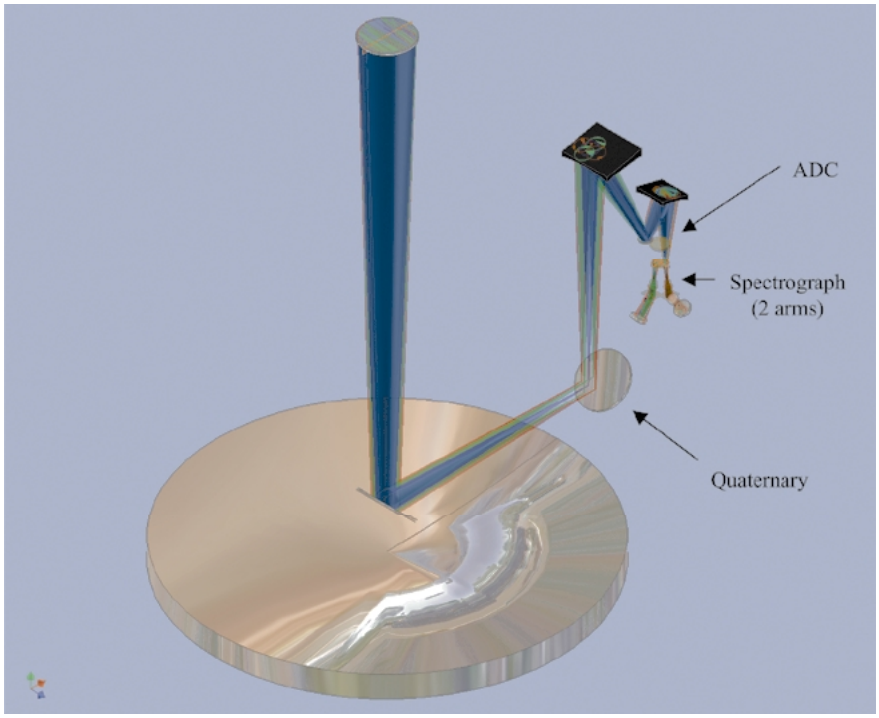


Figure 10.5. The concept for Damien Jones's 3MfR triple anastigmatic f/15 to f/5 relay, plus ADC is shown. The first mirror of the 3MfR is ~20 m from the vertical fold. Structural analysis may argue for the beam to be diverted downward with the system mounted in a cage below the Nasmyth platform.

The spectrograph: There are a variety of possibilities for the spectrograph design. However, the requirement to produce spectra over such a wide field falls easily into a VIMOS-like concept (Web reference 14) whereby the field is divided by four field lenses into contiguous quadrants. Each field-lens/collimator combination then creates its own pupil where the light is dispersed into four separate cameras. Such a format leads to individual FOVs of ~10 x 7 arcmin for each spectrograph (a total rectangular field of 20 x 14 arcmin with some vignetting losses in the extreme corners), which is imaged at ~f/1.1 onto a $6k^2$ (15-μm pixel) detector. At such speeds the cameras are, of necessity, internal focus Schmidt cameras (or their variants), but for such a beam size the vignetting of the detector package is < 10%.

A 400 mm collimated beam, combined with a mosaic of 1200 l/mm standard surface-ruled gratings mounted at a ~40° camera/collimator angle, would then produce first-order (λ~600 nm) resolutions R ~3600 for slit-widths of 0.75 arcsec over a ~120 nm λ-range. A full octave could be recovered at R~850. While VIMOS (and DEIMOS) use classical surface-ruled gratings, we suggest the use of volume phase holographic (VPH) gratings in a transmission Littrow configuration (Robertson, et al., 2000) to give higher dispersions and greater blaze efficiencies. As a comparison, a 2400 l/mm VPH so configured could give R~7600. While the necessity for articulated cameras creates more mechanical complexity, the efficiency and spectral resolution gains are substantial. Furthermore, camera articulation has the added benefit of permitting direct imaging by simple removal of the grating leading to a 0.1 arcsec sampled, 20 x 14 arcmin imaging field on CELT.

Full system V-band spectroscopic efficiencies of ~27% are anticipated, assuming new state-of-the-art LLNL reflective coatings for the 3MfR. Hence for the Poisson-noise limited (N&S) case where 50% of the exposure is on sky, the point-source 5σ magnitude limits (assuming 20% slit-jaw losses) at R~1000 are B~26.2 [R~25.0] in a 10^4 sec integration.

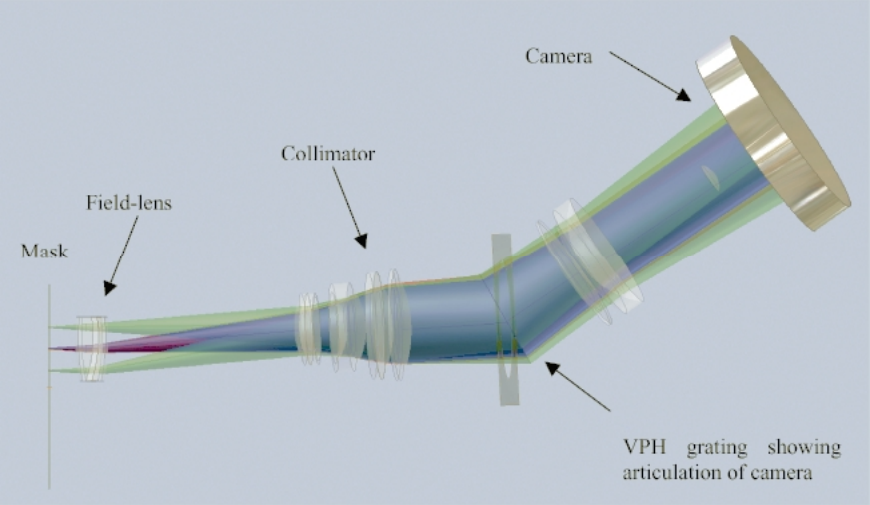


Figure 10-6. Optical layout of one of four arms of CFOS from the mask and field lens through collimator, VPH grating and articulated camera, to the detector.

10.5.2 Single-Conjugate AO Instrumentation

Mid-infrared camera/spectrograph

Basic requirements for this instrument are:

- wavelength range: 8-27 μm
- cold pupil for rejection of out-of-beam thermal background
- coronagraphic mode with occulting disk and rotating Lyot stop
- grism spectroscopy
- Ge grism to R ~ 1000
- Ge R4 echelle to R ~ 4000
- telescope focus located within dewar for cryogenic slits and occulting spot
- Nyquist sampling: 0.0344 arcsec/pixel at 10 μm; 0.0687 arcsec /pixel at 20 μm
- final F-ratio: F/5.4 (demagnification = 2.77)
- FOV: 70.4 x 70.4 arcsec (154 x 154 mm at F/15) with 2k x 2k As:Si array

The broad wavelength range and high throughput suggests an all-reflecting, non-obscured optical design. The f/15 collimator will have a FOV ~3.8˚ while the f/5.4 camera has a FOV ~10.6˚. This combination of f-number and FOV implies that simple configurations such as a pair of off-axis parabolas (OAPs) will not provide adequate control of aberrations for diffraction-limited performance. Both optical elements probably need to be TMAs, but a detailed design is not yet available.

The best detectors for this range are Si:As BIB devices. It is assumed that large formats will be developed for NGST and other applications. In estimating performance and design, the following detector properties were assumed: a $2k^2$ As:Si detector with a pixel pitch of 27 μm and no gaps (total detector size of 55.3

x 55.3 mm) and an operating temperature of 8° K. The dark current is assumed to be 1 e-/s, the readout noise is $4e^-$ rms, and the charge storage capacity or well depth is $3.10^7 e^-$.

For a pupil diameter of 10 mm we expect a collimator diameter of 164 mm, and the total optical path from cold focal plane to detector array should be about 410 mm. Cryogenic requirements at 8° K can be estimated from the volume and surface area of the dewar, which are ~0.0085m^3 and ~0.21m^2 respectively. Radiative heat loading power is ~ 70/n W, where n = number of insulating layers (typically three). The dewar and internal optical bench dimensions are likely to be a cylinder about 0.45 m diameter and 0.6 m long with a mass of about 170 kg.

The point-source sensitivity for broadband imaging at 10 μm (R ~10) should be about 30 μJy, 10σ, t =10^4s, assuming a 5% mirror emissivity. Control of diffraction and scattering from intersegment gaps in the CELT primary will be essential for high contrast imaging. Ideally, a two-mirror Cassegrain telescope optical system is preferred over a three-mirror Nasmyth, and a cryogenic AO system is also required.

10.5.3 Multi-Conjugate AO Instrumentation

Near-infrared Camera

The basic requirements for a near-IR AO camera and the implications of those requirements are as follows:

- Nyquist sampled PSF at 1 and 2 μm: implies 3.4 mas/pixel and 6.8 mas/pixel respectively
- practical MCAO FOV = 60 arcsec diameter: implies physical FOV = 131 mm
- wavelength range: 0.9-2.5 μm: HgCdTe detectors and "warm" AO system

Assuming the plate scale of the CELT (f/15) telescope = 0.458 arcsec/mm, and that the detector has 18 μm pixels, then two pixels is 8.25 mas. This is too coarse and therefore a magnification system is needed. Because of the wavelength range involved, it is best to have two camera systems. To achieve Nyquist sampling at both 1 and 2 μm the final camera focal ratios should be:

- FR_{cam} (1 μm) = 36, magnification = 2.4
- FR_{cam} (2 μm) = 18, magnification = 1.2

Adopting a conservative detector format of 4k x 4k pixels for now. Then the FOV is

- 4096 elements ==> 13.926 arcsec with 0.0034 arcsec per pixel at 1 μm
- 2048 elements ==> 13.926 arcsec with 0.0068 arcsec per pixel at 2 μm

To incorporate these requirements into a design and increase its efficiency, a four-shooter concept with twin-channel cameras is proposed. Figure10-7 shows one possible way to increase the fraction of FOV used, given the assumption that the largest arrays available will be $4k^2$ pixels. Assuming that the 4k x 4k detectors are made by butting 2k x 2k arrays, then each camera has five 2k x 2k detectors, giving a total of twenty 2k x 2k arrays. There need to be four identical electronic controllers, each capable of handling five arrays. In a more detailed design study, it would be attractive to investigate how to accommodate an 8k x 8k array.

If the thermal background is noticeable, then the Lyot stops must rotate as the pupil rotates. Without this complication, there would be only two mechanisms (filter wheels) per camera. See Figure 10-8 below which shows one of four cameras.

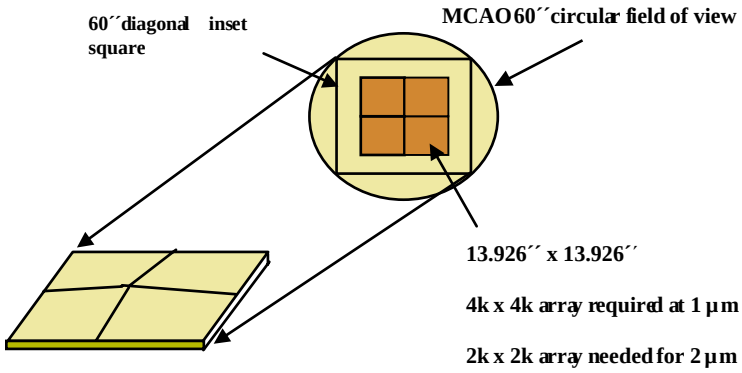


Figure 10-7. Four cameras are required, each with a ~14 arcsec FOV, to cover the central ~30 arcsec of the corrected field. To further increase efficiency, each camera could be twin-channel using a dichroic beam-splitter to send J and H to one beam and K to the other.

Packaging of the instrument could be accomplished within an enclosure of volume ~1 m cube with an up-looking entrance flange. The back focal length required from the flange face to the focal plane at the pyramid mirror would be about 0.25 m. Only liquid nitrogen temperatures are required. Each camera module could be equipped with an LN2 dewar or each camera could be cooled with a closed-cycle refrigerator. The total weight of the instrument would be of order 1 ton.

A cost analysis is given below, based largely on the detectors and previous experience with simpler cameras:

- four cameras x 5 arrays (each 2k x 2k @ $0.25M): $5.0M
- four sets of optics at $250k per set: $1M
- four dewars at $125k each: $0.5M
- four sets of electronics, each runs 5 arrays; @ $500k per set: $2M
- labor costs (design, assemble and test, software) ~ 100 myr @ $100k/myr: $10M

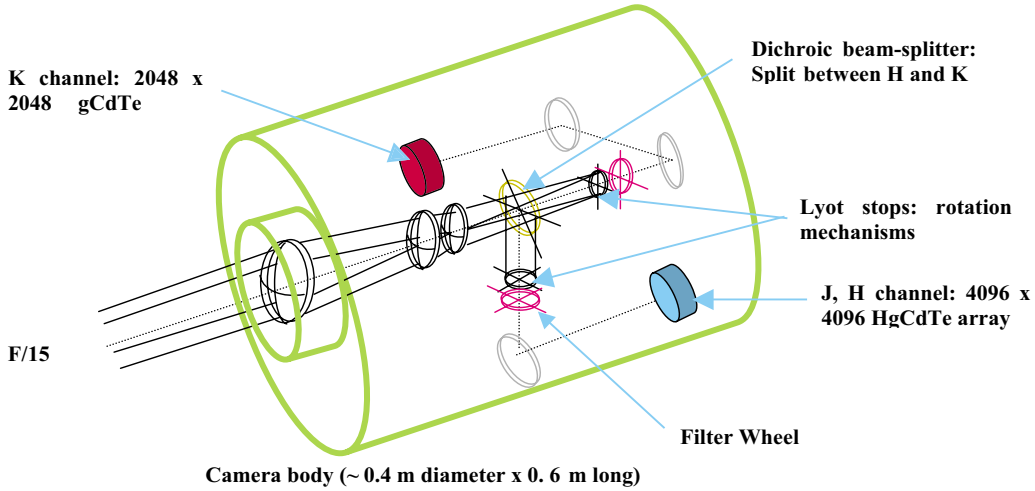


Figure 10-8. Four mechanisms: two filter wheels and two Lyot stops

Thus the total cost of the MCAO camera will be ~$18.5M, which is several times more expensive than, for example, the AO camera on Keck II (NIRC2).

Near-infrared fiber-based d-IFU spectrograph

We assume here an MCAO-fed d-IFU spectrograph having 20-50 sampling modules with 10 x 10 spatial elements sampling in the range 25-100 mas. Spectral resolutions R~4000 are assumed so that the full J, H, K windows are well matched to a 2k detector format.

We describe the use of multimode fibers with a core size of 50 μm and an outer fiber diameter of < 150 μm. While smaller (~35 μm) fibers would be better matched to the 50 mas spatial resolution, the 50 μm core fibers have become a standard for fine IFU systems, are mechanically robust, and reduce diffraction losses at their input. To avoid input filling-factor losses, each 10 x 10 fiber bundle has a matched lenslet array that forms individual pupil images at each fiber input. Each fiber/lenslet array module is then placed at a magnified telescope focus re-formed by small individual optics to give the correct 50 mas sampling scale.

A simple linear fiber slit will have poor output fill factor due to the fibers' core:clad ratio. However, if the fibers are staggered in the l-direction so that their interspectral distance is ~80 μm, this will increase the fiber packing density on the detector. If we assume a $2k^2$ (18 μm pixel) detector, then the spectrograph must map the 80 μm fiber spacing onto a 36 μm detector pitch to allow ~1000 fibers (10 d-IFU modules of 100 fibers each), thus requiring a collimator-camera ratio of ~2.2 (e.g., f/5 to f/2.2) with modest field angles.

The fiber-based d-IFU concept thus consists of a fore-optic region where a common cold pupil is formed. Multiple fiber bundles are positioned around this field; each bundle has a small magnifying lens and a lenslet array to sample and couple the light into the fibers. The cryogenic fibers are then brought to a set of identical spectrograph modules, each supporting 10 d-IFUs, the full multiplex being defined by the number of spectrographs and the details of the deployment scheme.

Several variations of the fiber concept are possible, including some of the spectrographic modules having different gratings settings. Also, a significant number of fibers (perhaps 1024) could be placed into a central bundle that does not move. This bundle would be most useful for those science cases where the multi-object capability is not warranted. In a cluster of galaxies, perhaps this central bundle could be used on the central cD galaxy, while the 70 other tentacles could explore other cluster members. It would also be possible to have different magnifying lenses in front of different bundles so multiple plate scales could be supported.

Near-infrared slicer-based d-IFU spectrograph

Introduction

We focus in this section on a possible CELT instrument concept that uses slicer d-IFUs for MCAO-fed near-IR spectroscopy (Taylor 2001d).

The baseline requirements of the d-IFU slicer system are summarized as follows:

- Multiple IFUs deployable over a 2 arcmin diameter, MCAO-delivered FOV. The term "multiple" here simply implies a number limited by either the chosen deployment scheme, the number of spectrographs, or the available detector real estate. We assume a reference of 16 d-IFUs; but this

figure is quite arbitrary and could certainly grow by factors of several should there be a sufficient scientific demand.

- To the extent possible, the user should be free to allocate d-IFUs to any multiple target field, although high surface density configurations may be impractical at some level to be defined.

- Given surface brightness of individual faint galaxy targets (Baldry and Taylor 2000), a rather conservative ~50 mas regime of spatial sampling is studied here. Options for both finer and coarser scales are available. However, we will not consider d-IFU systems designed for the ~7 mas J-band diffraction limit given the assumption that such work would be somewhat specialized in nature. Nevertheless, there is no reason why the basic concepts described here cannot be re-specified for finer scale work.

- The FOV of the IFU is limited by the spectrograph format and the number of d-IFUs. An arbitrary balance has been struck at 16 d-IFUs, each feeding 2k spectro-spatial X-pixels and 1k spatial Y-pixels. An arbitrarily rectangular 1.4 arcsec x 1.7 arcsec format has been chosen to match to the 50 mas spatial sampling (assuming an anamorphism of 2, to equalize 2-D spatial resolution). Clearly other aspect ratios can be defined; however, an inability to choose the position angle of an individual d-IFU might argue for square, or even circular, formats.

- We assume a wavelength range of 0.9-2.5 μm, covering the z-, J-, H- and K-bands defined by the sensitivity profile of Rockwell's Hawaii-2 arrays. These are chosen over the somewhat higher QE InSb arrays for their quieter read-noise and dark-count properties.

- A spectral resolution of R~4000 is required to effectively eliminate the OH sky lines in the data reduction process while retaining ~80% uncontaminated spectral samples. Rs > 4000 will not permit the full J-, H- or K-band to be recorded on the 2k available pixels, while Rs < 4000 will increase the fraction of OH-contaminated pixels (Taylor, et al., 1996).

- Given the significant investment in detector real estate (eight $2k^2$ HgCdTe Hawaii-2RGs) required for our 16 d-IFU reference system, it would seem prudent to investigate ways of using these for MCAO-fed imaging. In order to achieve near-diffraction-limited spatial resolution, a scale of ~4 mas/pixel has been chosen which leads to the prospect of 16 deployable imager units (or d-IUs) each having a ~4 arcsec x 8 arcsec FOV. This is presented as an alternative to a massive contiguous field imager that would require a pixel format of ~$32k^2$ (or a factor of 32 times the number of detectors) in order to fully sample the J-band diffraction limit over a 2-arcmin FOV.

Slicer deployment: Since the original 3-D image slicer development (Weitzel, et al., 1996), there have been several design iterations of the basic concept in order to facilitate a compact optical layout that will accommodate a number of different instrument configurations (Content 1997). Arguably the most mature of the current generation of slicers is that developed by the UK's ATC for the UIST instrument (Wells, et al., 2000). The performance of this device has already been verified in the laboratory environment; given its compact geometry, we will base the CELT d-IFU on its design as modified in the GIRMOS study (Wright, et al., 2000) (shown in Figure 10-9). Given the CELT design requirements itemized above, each d-IFU slicer module (with magnification and anamorphic reimaging relay) is likely to have dimensions of roughly 200 x 75 x 40 mm with the longer dimension parallel to the focal surface. The concept of simply deploying these slicer modules at the raw f/15 focus is clearly excluded

given the large ~90 arcsec footprint of the IFU; no more than four such units could be configured in a highly constrained fashion within the 2-arcmin MCAO FOV. We thus adopt the GIRMOS pick-off probe approach, modified to relay 16 chosen areas of the MCAO field into eight double slicer units, which themselves are arranged in pairs to feed four spectrographs.

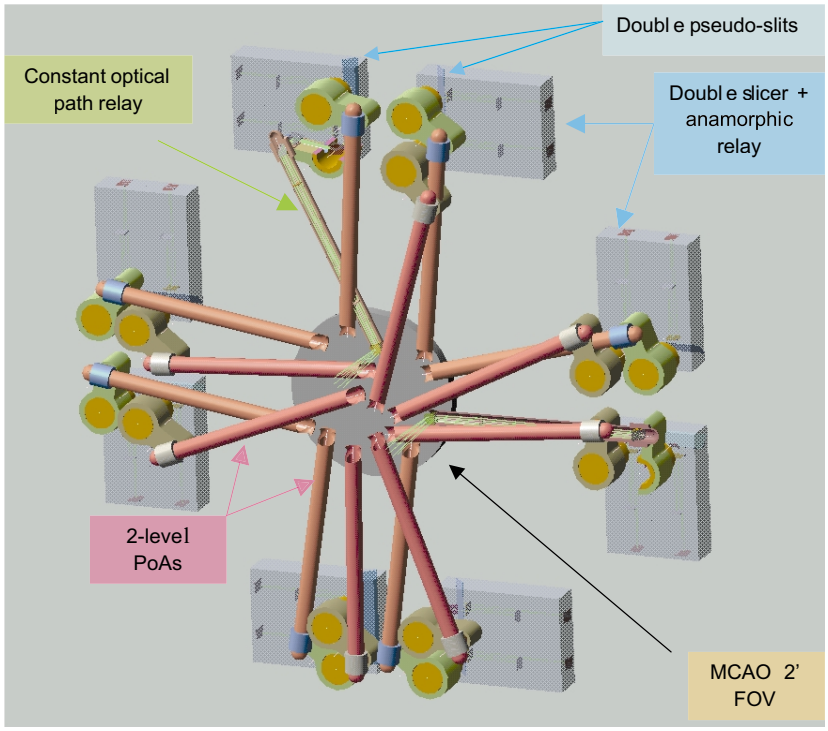


Figure 10-9. In front of the 2-arcmin circular FOV are 16 PoAs that relay each individual f/15 MCAO image into four pairs of double slicer units. Each PoA pair has different levels to permit crossovers within the field.

Details of the d-IFU optical train: There are three identifiable optical systems between the MCAO output focal-plane and the spectrograph input slit. The region of MCAO focal-plane selected by the pick-off arm (PoA) probe is first relayed through its PoA to a stationary fore-optics IFU feed. The latter then magnifies and anamorphizes the image as formed on the image slicer, while the slicer unit itself reformats the 2-D image formed by the fore-optics to create a pseudo-slit representing the spectrograph input. With such a complex optical train there is a large design degeneracy for each stage, with physical layout and mechanical engineering constraints defining optimal magnifications at each. We follow here the systems described in the GIRMOS study, since they are based on the experience of producing a real slicer for UIST, and on practical designs for the probes, slicers, and pseudo-slits. Naturally, the scales and magnifications for each stage are selected appropriately for the CELT case.

1. **The PoA probe:** The job of the PoA probe is to relay the native f/15 image through a cold stop (within the PoA) to a secondary ~f/30 image. Through the use of a selector mirror, this secondary, magnified image can be directed either to:
   i) the fore-optics of the slicer-IFU (d-IFU system), or
   ii) the re-imaging path of the d-IU system.

   Given the PoA's ~500 mm optical path length and limited ~25 mm input snout, a simple collimator/ camera relay is proposed to permit the larger field of the d-IU to pass unvignetted.

2. **The d-IFU anamorphic relay:** In entering the slicer-IFU, a three-mirror toroidal relay takes the secondary f/30 image and converts it into an ~f/140 image formed on the surface of the slicer, thus matching its 50 mas sampling to the ~1 mm physical thickness of each slice (chosen for ease of manufacture). The factor of ~2 induced anamorphism implies that the Y-direction of the image is formed at ~f/70.

3. **The slicer-IFU:** The 28 tilted slices of the slicer are each ~17 mm in length and are spherically profiled to form a linear array of 28 separate pupil images whose center-to-center distance is chosen to be ~2.5 mm, in accordance with the UIST prototype. Each individual pseudo-slitlet is then relayed through pupil mirrors to form a linear stack of slitlets with a total height of ~74 mm.

The deployable integral field unit (d-IFU) spectrographs: Each of four spectrographs services four d-IFU PoAs (16 PoAs in total), with their four slicers arranged in double pairs so that two sets of output slits are seen by the spectrograph. Each ~148 mm-long double pseudo-slit then produces an f/20 beam exiting in the spectro-spatial dispersion (or X-) direction and an f/10 beam in the spatial (or Y-) direction. We choose a fully transmissive, grism-based spectrograph for conceptual simplicity, although a full optical and engineering design evaluation may well reveal a more compact or efficient arrangement. A collimator focal-length of ~2 m is assumed to give adequate R~4000 spectral resolution, while limiting the field-angles into the collimator. A camera f-ratio of f/5 (X-direction) will thus produce an image of the pseudo-slit on the detector matched to 1k of its 18 μm pixels. In this way an arrangement of four d-IFU pseudo-slits can be supported by a single spectrograph having a 2k x 4k format formed from two close butted Hawaii-2RG arrays.

The relatively slow collimators and cameras would have modest field-angles and hence their optical design should not be too challenging. Each spectrograph is likely to be over 3.5 m in length, which may well argue for additional folds in the collimated beam, as was suggested in the GIRMOS study. Without such folds the four spectrographs could be arranged 90° apart in a ~1 m radius ring, and hence the whole d-IFU instrument from PoA to spectrograph would be contained in a ~3.5 m long, ~2.5 m diameter cylindrical dewar.

The deployable imager units (d-IU) system: Through the use of a selector mirror, the secondary f/30 image formed at the output of each PoA can bypass the slicer-IFU unit and be relayed, via a selectable mirror, directly onto the detector itself. A scale of 4 mas per pixel implies a 1:1 reimager incorporating an intermediate cold stop for the placement of suitable filters. Again, four PoAs would be able to service a single 4k x 2k array of two Hawaii-2RGs, giving each a 4 arcsec x 8 arcsec FOV and hence, all 16 PoAs could be used for non-contiguous MCAO imaging.

10.6 Space, Weight, and Deployment Considerations

Two large (15m x 30m) Nasmyth decks are available to support the instrumentation, and given the natural split between AO (near- and mid-IR) and seeing-limited (optical) instruments it would seem sensible to dedicate one platform to each. Furthermore, the CELT telescope design allows the tertiary mirror to articulate (in 2-D), permitting full use of the Nasmyth platform without the need to move each instrument into line with the altitude axis. The articulated tertiary then gives some considerable freedom in instrument location.

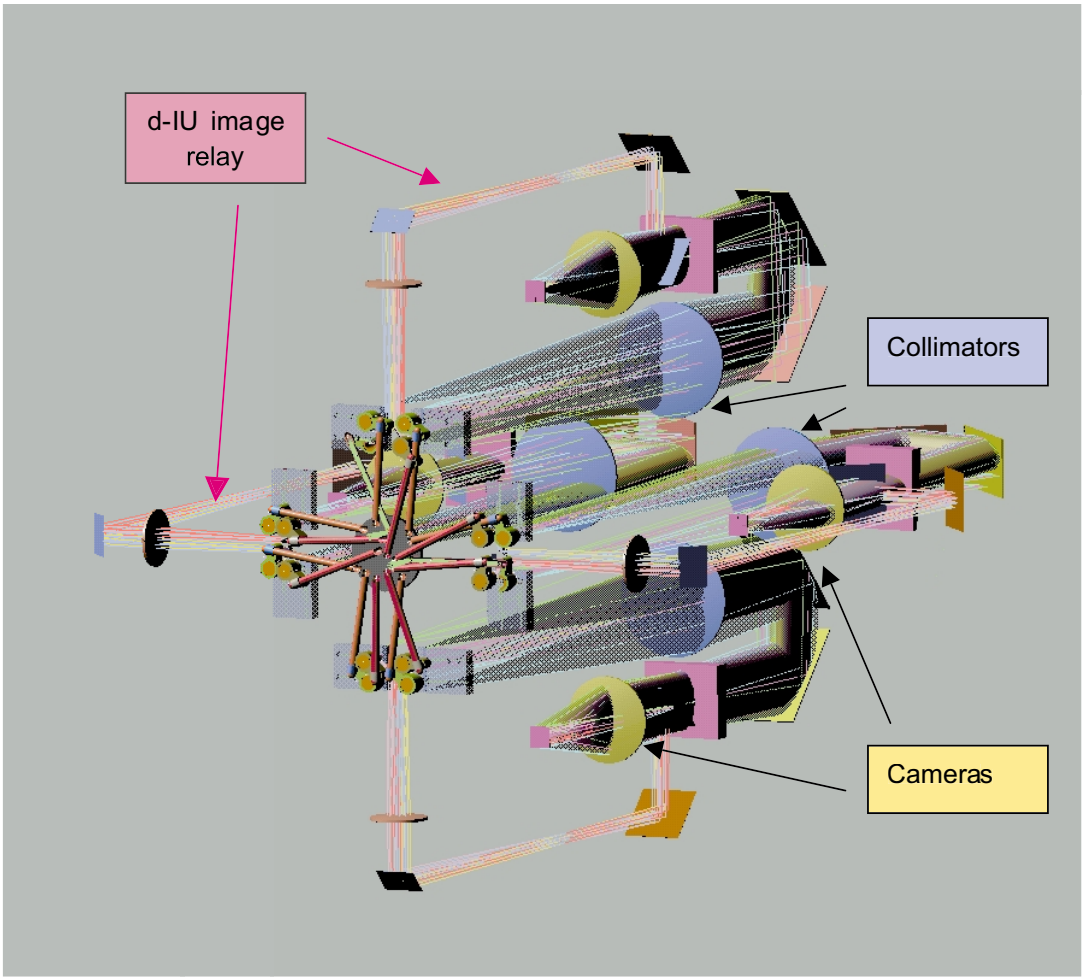


Figure 10-10. The full d-IFU system of 16 PoAs, eight double slicers and four spectrographs plus four d-IU relays is shown. Details of the subcomponents are given elsewhere (Taylor 2001d).

Seeing-limited (SL) Nasmyth: Figure 10-11 is a perspective drawing showing both (seeing-limited and AO) Nasmyth platforms populated by the fiber positioner (CfP) mounted on top of the MTHR spectrograph, whose space envelope is shown in yellow. We have shown the CFOS focal reducer/ ADC/spectrograph space envelope in its vertical mounting, while acknowledging that height (~20 m) constraints may force it to a different configuration.

In considering how instruments should take care of slow guiding (~2Hz tip-tilt and focus) on the seeing-limited Nasmyth platform, an important aspect of the CELT preliminary design will be to address how this is best done. In particular, with a platform dynamically decoupled (at some level) from the telescope itself, the question arises as to which part of the telescope optical train should be servoed. Initial thoughts would suggest that the tertiary mounting would have higher resonant frequencies than the secondary, and the fact that it is already articulated might suggest that this is a better optic to perform the ~2Hz compensation.

Adaptive Optics (AO) Nasmyth: Figure 10-11 also demonstrates a possible layout for the AO-Cam, D-IFU and GLAO spectrographs (CIRMOS). We have yet to include the space envelope for the AO system itself (supplying MCAO, GLAO and possibly LOAO). Issues concerning slow guiding for the AO Nasmyth are not as severe given the possibility of servoing one of the small internal mirrors in the AO chain.

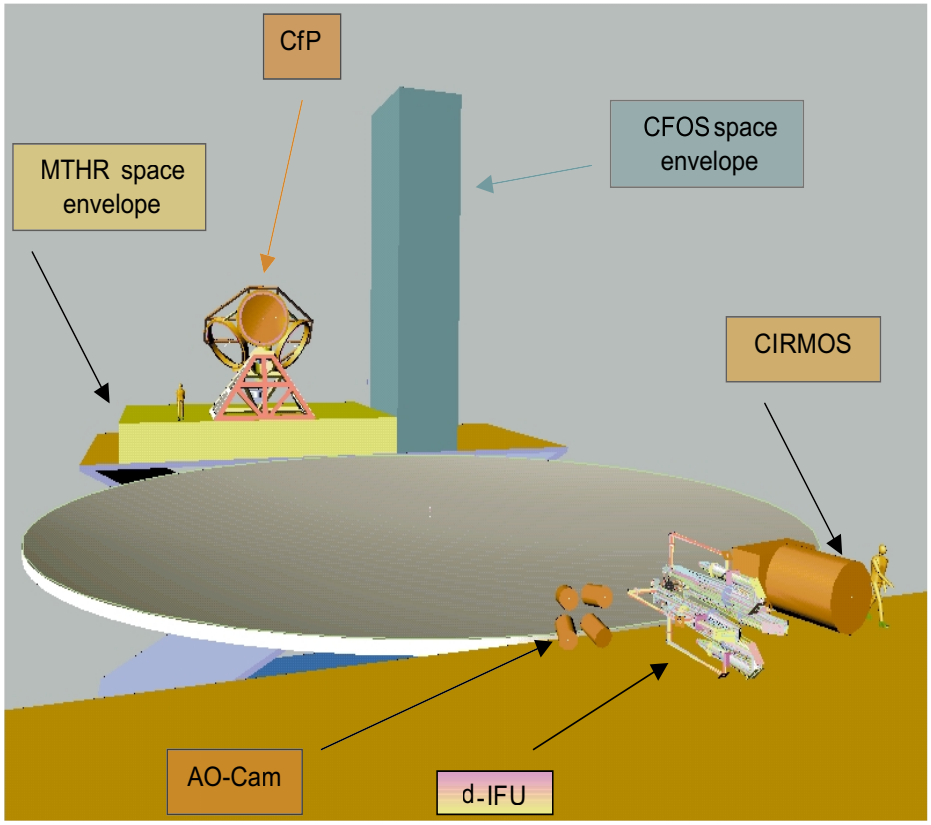


Figure 10-11. Layout of instruments on the SL and AO Nasmyth platforms. Very approximate space and weight budgets are given in Table 10.3.

Table 10-3. Space and weight budgets per instrument

Instrument	Dimensions (m) (Width*Depth*Height)	Weight (metric tonnes)
MTHR	16 * 10 * 2	63
CfP	6.5 * 5.5 * 8	30
CFOS	3.5 * 7 * 20	25
AO-CAM	1 * 1 * 1	1
D-IFU	2.5 * 2.5 * 3.5	4
CIRMOS	3 * 1.5 * 3	8

10.7 Cost Breakdown

Without detailed design studies, any cost estimates given for instrumentation are necessarily preliminary. Where possible, costs are based on materials and hardware, and on scaling up staff costs from existing (or planned) instruments.

Instrument	**Cost Estimate ($M)**
CFOS	30
CfP	8
MTHR	49
AOCAM	19
DIFU	23
MIRCAM/SPEC	8
TOTAL	**$137M**

Clearly, the total instrument budget required to capitalize on the telescope development and to produce real scientific returns in the first five years of operation will have to approach $100M. About 10% of this will need to be spent on an active research and development program, and in design studies during the next design phase for the telescope and dome; otherwise the instruments will simply not be ready for deployment. Many of these instruments are beyond the scope of any particular university research lab. It will be essential to establish a consortium of institutes and to subcontract to industry. Experience shows that initial concept studies, together with final system integration and testing by astronomers, is the best way to define, commission, and verify the instruments. Therefore, significant funding will be needed to support the university labs that generate the original concept.

In particular, such a necessarily huge investment in instrumentation demands that we actively seek ways to minimize costs without sacrificing functionality and performance. This *innovation factor* is targeted towards reducing costs, but requires that we invest early in a research and development program that specifically targets leading edge technologies necessary for the instrument suite. The following group of topics is geared towards generic technologies that span a range of instrument capabilities; bracketed acronyms refer to the instruments identified in Table 10-3.

- perfection of mirror slicer technologies through design, fabrication, deployment and test (d-IFU, MTHR)
- prototype and testing of cryogenic actuators for AO-fed instrumentation (d-IFU, CIRMOS, mid-IR)
- further development of VPH technologies and their cryogenic evaluation (CMOS, d-IFU, CIRMOS)
- minimization of scattered light in near-IR instrumentation (OH-suppression)
- large grating mosaics (MTHR, CMOS).
- development of high R (R5 to R10) gratings in the near-IR and optical (MTHR, mid-IR)
- large format VPH grating development (CFOS, MTHR, d-IFU)
- cryogenic chamber performance for 3-meter sized instruments (d-IFU, CIRMOS, mid-IR)
- engineering studies of large optic fabrication and mounting (ALL)
- detector development (ALL)

The following list is particular to individual instruments.

- perfection of fiber IFUs/slicer technologies through design, fabrication, deployment and test (MTHR)
- further opto-mechanical design concepts for seeing-limited multi-slit spectrographs (CFOS)
- further opto-mechanical design concepts for seeing limited focal-reducer and ADCs to feed multi-slit spectrographs (CFOS)
- a study of modal noise behavior in fibers to verify if it is a potential limitation for 1 part in 1000 sky subtraction (MTHR)

In the next phase of the CELT project we plan to review the science case in light of our currently identified instrument solutions and identify the capabilities required for first-light instrumentation on the seeing-limited and AO-foci. The output of this process will not only be a decision on a first-light instrumentation suite but will also prioritize and amend the current list of research and development activities, gearing it to optimizing performance. It will also minimize project risk and cost through planned-for innovation factor gains. This strategy summarizes the approach to delivering a CELT instrumentation package that is both focused on the overall science case and is practically achievable within a budget commensurate with the scale of the overall CELT program.

REFERENCES

Atherton and Taylor. 1982. "TAURUS, A Wide-Field Imaging Fabry-Perot Spectrometer for Astronomy." Monthly Notes of the Royal Astronomical Society **201**, 661.

Baldry and Taylor. 2000. "Design Study for MCAO Deployable IFUs: Sensitivities and Source Structure." <http://www.gemini.edu/sciops/instruments/adaptiveOptics/MCAOWorkshop_agenda.html>

Baldry and Taylor. 2001. "System Modeling" CELT Technical Note No. 12.

Barden, et al. 1993. "Hydra-Kitt Peak Multi-Object Spectroscopic System." Fiber Optics in Astronomy II. ASP Conference Series **37**, 185.

Bigelow and Nelson. 1998. "Determinate space-frame structure for the Keck II Echellette Spe ctrograph and Imager (ESI)." SPIE Proceedings **3355**

Bland-Hawthorn. 1996. *Anglo-Australian Newsletters* **79**, 14

Bland-Hawthorn and Jones. 1997. " TTF: a flexible approach to narrowband imaging." astro-ph/9707315. (See also http://www.aao.gov.au/astro/ttf.html).

Content. 1997. "A new design for integral field spectroscopy with 8-m telescopes." SPIE Proceedings **2871**, 1295.

de Bruijne, et al. 2001. "Analysis of Astronomical Date from Optical STJs." The First Galway Workshop on High Time Resolution Astrophysics. ASP Conference Series (in press).

Estrada, et al. 1998. "Large Format IR Arrays for Future Space and Ground-Based Astronomy Applications." SPIE Proceedings, Infrared Astronomical Instrumentation **3354**, 11.

Gillingham, et. al. 2000. “Echidna: A Multifibre Positioner for the Suburu Prime Focus.” SPIE Proceedings **4008**, 1395

Glazebrook and Bland-Hawthorn. 2001. “Microslit Nod-Shuffle Spectroscopy.” *Publication of the Astronomical Society of the Pacific* **113**, 197

Haynes, et al. 2000. “Characterization of cooled infrared fibers for the Gemini IRMOS.” SPIE Proceedings **4008,** 1203.

Hill, et al. 1986. “Deployment of the MX Spectrometer.” SPIE Proceedings **627**, 303. See also http://abell.as.arizona.edu/~hill/mxspec/

Jones and Taylor. 2001. “2MfR.” CELT Technical Note No. 19.

Kenworthy, Parry and Taylor. 2001. “SPIRAL Phase A: A Prototype Integral Field Spectrograph for the AAT.” *Publication of the Astronomical Society of the Pacific* **113**, 215

MacKenty, et al. 2000. “A Multi-Object Spectrometer using Micro-mirror Arrays.” Next General Space Telescope MOS study team.

McLean, et al. 1998. “Tehe Design and Development of NIRSPEC: A Near-Infrared Echelle Spectrograph for the Keck II Telescope.” SPIE Proceedings **3354**, 566

Mosely, et al. 2001. “Status of the development of a 128x128 microshutter array.” SPIE Proceedings **4178**

Oke, et al. 1994. “The Low Resolution Imaging Spectromter for the Keck Telescope.” SPIE Proceedings **2198**, 178

Pasquini, et al. 2001. “FLAMES: a multi-object fiber facility for the VLT.” SPIE Proceedings **4008**, 129

Rockwell Science Center. 2002. “Preliminary draft input on infrared sensors.” CELT Technical Note No. 25.

Robertson, et al. 2000. “ATLAS: A Cassegrain Spectrograph Based on Volume Phase Holographic Gratings.” SPIE Proceedings **4008**, 194

Taylor, et al. 1996. “The AUSTRALIS Study Report.” <http://msowww.anu.edu.au/~colless/AUSTRALIS/>

Taylor, et al. 1997. “The Anglo-Australian Observatory 2dF Project.” SPIE Proceedings **2871**, 145

Taylor and Gray. 1990. “Design Study for a New Wide-Field AAT Prime-Focus: The 2dF.” SPIE Proceedings **1236**, 290.

Taylor and Gray. 1994. “2dF: The AAT’s Planned Wide-Field Multi-Fibre Spectroscopic Survey Facility Report on Commissioning the 2dF Corrector/ADC.” SPIE Proceedings **2198**, 136

Taylor. 2001a. “Large Aperture Fast Cameras.” CELT Technical Note No. 13.

Taylor. 2001b. “Fibers for CELT.” CELT Technical Note No. 15.

Taylor. 2001c. “Fiber Aperture Choices for MTHR.” CELT Technical Note No.16.

Taylor. 2001d. “Slicer-based d-IFU Spectrograph.” CELT Technical Note No.18.

Taylor. 2001e. “IFUs for Palomar.” CELT Technical Note No. 21.

Vogt, et al. 1994. “HIRES: the High Resolution Echelle spectrometer on the keck Ten-Meter Telescope.” SPIE Proceedings **2198**, 362

Vogt, Steve. 2001. “CHRS and MTHR.” CELT Technical Note No. 20.

Weitzel, et al. 1996. “3D: The Next Generation Near-Infrared Imaging Spectrometer.” Astromony and Astrophysics Supplement Series **119**, 531

Wells, et al., 2000. “Design and testing of a cryogenic image slicing IFU for UKIRT and NGST.” SPIE Proceedings **4008**, 1215.

Wright, et al. 2000. “GIRMOS: an infrared multi-object spectrograph for Gemini.” SPIE Proceedings **4008**

WEB REFERENCES:

1. http://www.keck.hawaii.edu/realpublic/asteroid
2. http://www.astro.ufl.edu/~elston/flamingos/flamingos.html
3. http://www.hq.eso.org/instruments/fors1/index.html
4. http://www.aao.gov.au/2df/index.html
5. http://www.aao.gov.au/local/www/ozpoz
6. http://www-obs.univ-lyon1.fr/tiger/home_tiger.html
7. http://www.astro.ucla.edu/~irlab/osiris/
8. http://www.mpe-garching.mpg.de/www_ir/ir_instruments/3D/references
9. http://www.cfht.hawaii.edu/Reference/bulletin/Bull32/bull32I.html
10. http://www-phys.llnl.gov/Research/NGST/
11. http://www.aao.gov.au/astro/spiral.html
12. http://www.eso.org/instruments/flames
13. http://www.aao.gov.au/2df/index.html
14. http:www.hq.eso.org/instruments/vimos/index.html

Chapter 11. Expected Wavefront and Image Quality

11.1 Introduction

Error budgets serve multiple purposes. They provide a statistical estimate of the expected performance. They serve as a management tool for making trades between performance, budget, and schedule. In addition, they are the basis for developing tolerances. Since they will be used by a variety of people for these multiple purposes, they need to be practical tools. We have selected the form and language of the budgets so they can be practical tools, and so that errors can be readily calculated and combined. In order to achieve this we have necessarily made compromises and approximations.

In the budgets developed to date we have emphasized the central region of the point spread function. Diffraction effects will have significant impact on the outer regions of the point spread function and are important for extreme adaptive optics systems that will be used to image planets around nearby stars. The error budgets for that particular AO system and application need to be developed separately. Diffraction effects are shown in Tables 4-5 to 4-8.

Currently we have only considered the budgets for visible light. The important and additional requirements for infrared observing will be developed in the next design phase.

11.2 Error Budgets

CELT Report No. 10 (Mast and Nelson 2001) describes the CELT error budgets in detail. The motivation for the general form of the budgets, the budgets themselves, and extensive calculations are given to support the values in the budgets. The calculations cover many more pages than can be included here.

These error budgets are preliminary. They are based on a telescope optical design that is not final. Through the remainder of the CELT project the error budgets will evolve as this optical design evolves, the performance prediction calculations are refined, and cost tradeoffs are made. To allow changes to be more readily incorporated, where possible we have used analytic expressions and spreadsheets to describe the errors.

CELT will operate in two modes: Adaptive-Optics-Off and Adaptive-Optics-On. In Sections 11.3 and 11.4 we present error budgets for each mode. We have followed the general approach that was successfully used for the Keck telescopes and is described in Keck Observatory Reports (Nelson, Mast, and Faber 1985; Mast and Nelson 1986a, 1986b, 1987; Rockey 1986).

The telescope will be created in stages: design, fabrication, assembly, warping harness adjustment, optics alignment, open-loop active control, closed-loop active control, guiding, and adaptive optics. Budgets are required to define the tolerances required for each stage. The error budgets developed to date are only the final performance budgets. These are the residual errors expected after the above processes are completed and operating properly. They describe the final image or wavefront quality.

Observing Variables (z, θ, λ)
Errors budgets are in principle a function of zenith angle (z), field angle (θ), and wavelength (λ).

Zenith Angle (z): We have written the error budgets for the zenith, $z = 0$. The goals are most restrictive at $z = 0$, since atmospheric seeing degrades with increasing z. A full zenith-angle dependent error budget may be required. For the Keck telescopes the budget was described in terms of coefficients for zenith angle functions: constant, sin(z), and cos(z) (Mast and Nelson, 1986a). Our goals and budgets

are sometimes expressed in terms of image size and sometimes in terms of rms wavefront errors. For images dominated by atmospheric-turbulence

$$\text{image size} \sim (\sec z)^{3/5} \qquad \text{rms wavefront} \sim (\sec z)^{1/2} \tag{11-1}$$

Field Angle (θ): We will write the error budgets for on-axis observing. The only telescope error that increases with field angle is the optics design error, Ritchey-Chrétien astigmatism (see Section 4.2.1). Unlike the telescope errors, the adaptive optics system field angle dependence is complicated and depends on the particular adaptive optics system mode.

Wavelength (λ): To an excellent approximation the atmospheric-turbulence-induced wavefront errors are independent of wavelength. For index of refraction $n(\lambda)$ the variation with wavelength, λ, is $[(n(0.5)-n(2.0)]/[n(0.5)-1.0] = 0.022$. They are most clearly described as rms pathlength errors in units of meters (σ). The wavelength of observation becomes relevant when we convert the wavefront errors to image parameters. Then, the image FWHM and Strehl ratio legitimately depend on wavelength.

$$\text{FWHM} \sim \lambda^{-1/5} \qquad \text{Strehl ratio} = \exp\left(-(2\pi\sigma/\lambda)^2\right) \tag{11-2}$$

External Drivers

For each of atmospheric seeing, wind, and temperature there are probability distributions that depend on the site and the dome design. For each probability distribution we will select a percentile cut to be used to specify the telescope design and error budgets. For example, we might specify that the error budgets must be achieved under conditions specified by a best seeing value, a maximum wind velocity, a maximum temperature range, and a maximum temperature gradient.

Atmospheric Seeing: Although a site has not been selected for CELT, for the initial budgets we make a baseline assumption about atmospheric seeing. We assume that for only 10% of the time at the CELT site the FWHM seeing will be better than 0.30 arcsec ($\lambda = 0.5$ μm). Once a site has been selected this 10% value may be modified.

Wind: We are currently collecting wind data from candidate sites and studying the effects of wind on the dome, the telescope top end, and the telescope primary mirror. Thus the wind conditions under which the error budgets must apply have not yet been established.

Temperature: We are currently collecting temperature data from candidate sites and studying the ability to control the temperature variation seen by the telescope. Until that analysis for a selected site is available we assume the operating temperature range is 2 ± 4 °C (the 90 percentile range for Mauna Kea).

Merit Functions and Parameters (f/p)

We need to express the goals and the budgeted performance in a precise, engineering-based language. There are two general possible categories, wavefront errors and image blur. Within each category there are further choices, each with its advantages and disadvantages. A description of these considerations is given in CELT Report No. 10 (Mast and Nelson 2001).

We will write the budget for the AO-off mode in terms of the 80% enclosed energy diameter, θ(80), in arcseconds on the sky. This is an image blur figure of merit. We will write AO-on budget in terms of rms wavefront errors (in units of nm) that are residual after AO correction.

For each choice above, a method for combining the error is also required. Methods differ in their computational ease and the intuitive clarity they allow. We will combine wavefront quality rms wavefront by adding the rms values in quadrature. We will combine 80% enclosed energy diameters by adding diameters in quadrature.

11.3 Adaptive-Optics-Off Error Budgets

The following two tables give the current budget for the telescope as a whole and a detailed budget for the primary mirror component. These are budgets for the adaptive optics system off, the Nasmyth focus, zenith angle = 0 degrees, operating temperature range 2±4 °C. They are written in terms of the image diameter in arcseconds that encloses 80% of the image energy from a point source. The right column gives the number of the supporting calculation in CELT Report 10.

We have assumed in this initial budget that the Telescope Guiding includes some of the tertiary alignment errors. Some of the errors are yet to be calculated, and the values here have been estimated from those in the Keck error budgets.

Table 11-1. Adaptive-optics-off error budget for the telescope

				CELT Rpt. 10 Calculation
zenith angle = 0 degrees, field angle = 0 degrees, wavelength = 0.5 microns				1-3
Image quality using 80% enclosed-energy diameter (arcsec on sky)				4
Atmosphere			0.567	5
Telescope			0.260	6
Optical Design		0.000		7
Primary		0.204		8
Secondary		0.100		
Surface Figuring	0.096			9
Support*	0.029			10
Tertiary		0.034		
Surface Figuring	0.032			11
Support*	0.011			12
Secondary alignment		0.084		
passive degrees of freedom (1)	0.000			13
active degrees of freedom (5)	0.084			
desired sensor readings	0.042			14
sensor noise	0.042			15
actuator noise	0.042			16
uncontrolled frequencies	0.042			17
Telescope Guiding and Tertiary alignment		0.088		23
	quadrature sum=	**0.260**	**0.624**	

*** Both secondary and tertiary are likely to have active control of their surfaces. In this summary error budget we assume the residual errors are part of support errors.**

Table 11-2. Adaptive-optics-off error budget for the primary mirror

				CELT Rpt 10 Calculation
zenith angle = 0 degrees, field angle = 0 degrees, wavelength = 0.5 microns				1-3
Image quality using 80% enclosed-energy diameter (arcsec on sky)				4
Primary Mirror			0.204	35
Segment Figuring			0.144	36
Focus		0.100		37
Astigmatism		0.050		38
Coma		0.058		39
Higher order		0.070		40
Segment Thermal Distortion			0.075	
Variations in mean CTE		0.000		41
Gradients in CTE		0.062		42
Gradients in temperature		0.043		43
Segment Support			0.091	
Axial design		0.064		44
Axial fabrication/assembly		0.008		45
Lateral design		0.000		46
Lateral fabrication/assembly		0.003		47
Thermal effects		0.032		48
Gravity effects		0.010		49
Segment Alignment - Passive			0.055	
Initial alignment		0.055		
radial	0.010			50
azimuthal	0.054			51
Temp-induced segment motion		0.005		52
Gravity-induced segment motion		0.020		53
Segment Alignment - Active			0.053	
Desired Displacement Sensor Readings		0.025		54
Desired Wavefront Sensor Readings		0.025		55
Displacement Sensor Readings		0.006		
thermal effects	0.010			56
gravity effects	0.010			57
temporal drift	0.010			58
sensor noise	0.010			59
Wavefront Sensor Centroids		0.000		
thermal effects	0.010			60
gravity effects	0.010			61
temporal drift	0.010			62
centroid noise	0.010			63
Actuators		0.023		64
Uncontrolled frequencies		0.031		
wind	0.030			65
seismic	0.006			66
equipment	0.006			67

An error budget here is only the telescope contribution. The AO system errors described in Chapter 9 must be added in quadrature to those given here.

The deformable mirrors of the AO system will correct some of the lower spatial frequency telescope errors. To be specific, we have assumed here that the deformable mirror has 3000 actuators. We have assumed that the secondary and tertiary alignment errors are well corrected by these.

11.4 Adaptive-Optics-On Error Budgets

The following two tables give the current budget for the telescope as a whole and a detailed budget for the primary mirror component. These are budgets for the adaptive optics system on, the Nasmyth focus, zenith angle = 0^o, the operating temperature range $2\pm 4\ ^oC$, and they are written in terms of the residual rms wavefront error.

Table 11-3. Adaptive-optics-on error budget for the telescope

				CELT Rpt 10 Calculation
zenith angle = 0 degrees, field angle = 0 degrees, wavelength = 0.5 microns				1-3
Wavefront quality using rms wavefront in nanometers				4
After Correction by the Deformable Mirror(s)				
N_{DM} actuators = 3000				
rms wavefront (nm)				
Telescope			50.2	6
Optical Design		0.0		7
Primary		45.0		8
Secondary		18.4		
Surface Figuring	17.5			9
Support*	5.8			10
Tertiary		11.3		
Surface Figuring	10.7			11
Support*	3.6			12
Secondary alignment		0.0		
passive degrees of freedom (1)	0.0			13
active degrees of freedom (5)	0.0			
desired sensor readings	0.0			14
sensor noise	0.0			15
actuator noise	0.0			16
uncontrolled frequencies	0.0			17
Telescope Guiding		5.0		23
	quadrature sum = 50.2			

*** Both the secondary and tertiary are likely to have active control of their surfaces. In this summary error budget, we assume the residual errors are part of support errors.**

An error budget here is only the telescope contribution. The AO system errors described in Chapter 9 must be added in quadrature to those given here.

The deformable mirrors of the AO system will also correct some of the lower spatial frequency segment errors. We have included this correction assuming 3000 deformable mirror actuators in the segment figuring errors. Mast and Nelson (2002) give a calculation of the correction of segment figuring errors versus deformable mirror actuator value. We assume off the low order aberrations measured by the Wavefront Control Sensor are corrected.

Table 11-4. Adaptive-optics-on error budget for the primary mirror

				CELT Rpt 10 Calculation
zenith angle = 0 degrees, field angle = 0 degrees, wavelength = 0.5 microns				1-3
Wavefront quality using rms wavefront in nanometers				4
N_{DM} actuators = 3000				
Primary Mirror			44.9	35
Segment Figuring		26.5		36
Focus	18.3			37
Astigmatism	12.8			38
Coma	9.5			39
Higher order	10.6			40
Segment Thermal Distortion		11.5		
Variations in mean CTE	1.7			41
Gradients in CTE	11.0			42
Gradients in temperature	2.9			43
Segment Support		17.7		
Axial design	5.1			44
Axial fabrication/assembly	13.0			45
Lateral design	3.0			46
Lateral fabrication/assembly	3.2			47
Thermal effects	10.0			48
Gravity effects	0.0			49
Segment Alignment - Passive		23.5		
Initial alignment	23.4			
radial	8.0			50
azimuthal	22.0			51
Temp-induced segment motion	2.0			52
Gravity-induced segment motion	0.0			53
Segment Alignment - Active		17.8		
Desired Displacement Sensor Readings	5.0			54
Desired Wavefront Sensor Readings	0.0			55
Displacement Sensor Readings	8.5			
thermal effects	2.0			56
gravity effects	0.0			57
temporal drift	0.0			58
sensor noise	8.3			59
Wavefront Sensor Centroids	0.0			

Table 11-4, continued...

thermal effects	0.0		60
gravity effects	0.0		61
temporal drift	0.0		62
centroid noise	0.0		63
Actuators		10.6	64
Uncontrolled frequencies		10.4	
wind	10.0		65
seismic	2.0		66
equipment	2.0		67

REFERENCES

Nelson, J., T. Mast, and S. Faber. 1985. “The Design of the Keck Observatory and Telescope.” Keck Observatory Report No. 90.

Mast, T., and J. Nelson. 1986a. “Error Budget for the Keck Observatory Primary Mirror.” Keck Observatory Report No. 144.

Mast, T., and J. Nelson. 1986b. “Image Quality Error Budgets For the Keck Telescope.” Keck Observatory Report No. 158.

Mast, T., and J. Nelson. 1987. “The Error Budget For The F/15 Secondary Mirror.” Keck Observatory Report No. 157.

Mast, T., and J. Nelson. 2002. “Deformable Mirror Fitting Errors for Zernike Aberrations in an Array of Hexagons.” CELT Technical Note No. 1.

Mast, T., and J. Nelson. 2001. “CELT Error Budgets.” CELT Report No. 10.

Rockey, D. 1986. “Absolute Pointing Error Budget.” Keck Observatory Report No. 159.

12. Observatory Dome and Facilities

12.1 Introduction

CELT facilities will be located at the summit, at local headquarters, and at remote observing facilities in California. An extensive description of the requirements, candidate designs, and all major issues for these facilities are given in CELT Report No. 23 (Tytler and Nelson 2002). We present here only the salient features.

First and foremost, these facilities will be designed to provide safety for personnel and equipment. They will also be designed to minimize the cost of operations. They must be designed to support the telescope over its expected life of greater than 50 years.

In Sections 12.2, 12.5, and 12.6 we describe the requirements for the facilities at each location. Section 12.3 describes a reference design concept for the enclosure while 12.4 describes related thermal issues. Section 12.7 outlines work to be carried out in the preliminary and final design phases. Foremost among these will be to study the candidate enclosure designs and to establish a reference design.

12.2 Summit Facility Requirements

At the summit there will be a telescope enclosure and an associated building to support the telescope operations. The enclosure may be in any one of various possible shapes: dome, box, cylinder, or hybrid. In Section 12.3 we focus on a spherical dome supported by a fixed cylindrical wall in order to give specific numbers.

12.2.1 Enclosure Requirements

Not all the enclosure requirements can be explicitly defined at this time. We give here the requirements categories and define those that are currently available.

Enclosure Size

The size of the telescope structure drives the size and cost of the enclosure. The critical dimension is the distance from the elevation axis to the top of the telescope. This is currently 38 m. This distance depends on the optical design and the support chosen for the secondary mirror. The vertical distance from the azimuth to elevation bearings influences the height of the center of the enclosure, and hence the height of a fixed cylindrical wall. The distance from the elevation axis to the back of the primary mirror support and the access required for maintenance may also impact the enclosure. The final optical and structural designs are required to specify the exact dimensions.

Slit Size

The enclosure must provide visual access to the sky over the circle of the astronomical opening, a circle with a 16.25 m radius (centered on the optical axis at a dome radius of 45 m) equal to an opening radius angle from the sphere center of 21.168°. The enclosure opening must be adjustable in position to allow a minimum opening (to provide maximal wind protection) that follows the telescope beam at all accessed elevations and azimuths. The minimum travel in zenith angle is 0-65°. All azimuth angles must be accessible. A fully open slit, which would allow the telescope unobstructed viewing, would extend 107.336° comprising the 65° motion of the telescope plus 21.168° at both the top and bottom.

Environmental Protection

Sun

The enclosure should reflect sunlight to reduce the daytime heat load, thus minimizing the refrigeration required to maintain the telescope and enclosure interior temperatures at the anticipated nighttime temperature.

Temperature

The enclosure must insulate the telescope from the outside temperature to minimize the refrigeration required to maintain the telescope and enclosure interior temperatures at the anticipated nighttime temperature.

Humidity

Moisture in the enclosure must not be allowed to condense on optical surfaces at any time. Condensed moisture "welds" dust to the surface and prevents effective cleaning by CO_2 spraying.

Rain

The enclosure must be sealed against all ordinary and wind-driven rain.

Snow and Ice

The enclosure must be able to safely carry the 100-year maximum snow and ice loads for the site, estimated as 20 kg/m^2 (for Mauna Kea). Enclosure shape should minimize snow and ice buildup, seals should be designed to prevent the buildup of ice and snow, and drainage channels are needed to remove water from the vicinity of seals and moving parts. Strip heaters may be needed to melt ice and snow near moving joints. The enclosure must prevent the possibility of snow or ice blowing or falling through the astronomical opening onto the telescope.

Wind Protection and Attenuation

We assume a specification defined by the wind conditions on Mauna Kea. The median wind speed is 7m/s, and 90% of the time it is under 12m/s. The telescope should be usable, up to 12m/s, in most sky directions, with some degradation of performance in some sky directions. The enclosure must survive winds up to 70 m/s. The top end of the telescope and the primary mirror must be shielded by the enclosure and shutters to prevent wind forces from degrading the image quality.

Vibration

All equipment (pumps, motors, drives, etc.) must be designed, controlled, and isolated to minimize vibrations being introduced into the telescope either through the ground or the air.

Earthquake

The enclosure must be designed to survive a 100-year earthquake.

Foundations

The foundations should minimize the transmission of wind-induced vibrations of the enclosure into the telescope.

Access to the Enclosure

Information, people, and equipment will need to travel frequently to and from the site, during construction and throughout the life of the telescope. Access will be affected by particulars of airplane routes from

California, the nature of the roads from the airport to the base site, and roads from the base site to the telescope.

Small equipment and roughly 40 people will travel daily from the base site to the telescope. Access roads must allow safe travel under conditions of driver fatigue, severe weather, and darkness. They should be designed to minimize wear and tear on drivers, passengers, and vehicles. Large equipment will travel frequently during construction and then a few times per year. During the final design stage equipment and component sizes will be determined, and these will allow us to specify the requirements for road width, height clearance, and maximum grade.

Access for Assembly and Maintenance

The enclosure must not impede access to and handling of the telescope components. The enclosure must provide access to the secondary mirror, shutters, the topmost point of the enclosure, bogies and drives, any ventilation doors, lighting fixtures, sensors, and cameras.

Communications

The following communications infrastructure will be required.

- High-speed, high-capacity and high-reliability voice, video, and data links will be essential between the telescope site, the base site, and California remote observing sites.
- Wireless devices will be used both on and off the various sites for voice, video, and data communications.
- Real-time communication links with local aircraft control and satellite protection agencies will be required to enhance AO laser system safety.

Motion

The enclosure opening to the sky (slit) must rotate at least 360° in azimuth to keep the astronomical opening aligned with the telescope at all times. The motion must be continuous and smooth to minimize vibrations.

Slit Shutters

A system of shutters must open the slit to the sky for observing and at the same time provide maximum protection from wind and dust. When closed it must seal to meet all the protection requirements listed above.

Lighting

Light sources are required to illuminate:

- the entire enclosure volume to approximately 100 lux (lumens/m^2)
- areas where people frequently move and work, including the doors, main passageways, elevator entries and exits, and the entire Nasmyth platform to 300 lux
- any specified part of the telescope or enclosure with spot lighting of 1000 lux (these sources should be wall mounted, and the pointing direction must be remotely controlled)
- places of potential collisions (of people, telescope, cranes, etc.), with strip lighting

Lighting sources should have the highest practical energy efficiencies to minimize heating and conserve energy.

Electrical Power and Noise
The major power requirements will be exhaust fans, lighting, air refrigeration, computers, and instruments. Additional power requirements include enclosure motion, shutter motion, and emergency power. Tytler and Nelson (2002) give estimates of the CELT requirements. They conclude that the mean power will be about 900 kW, with a peak requirement of about 2100 kW. They recommend the mountain facilities be rated for 3000 kW.

Uninterruptible and emergency power supplies will be required to maintain emergency lighting, computers, communications links, instrument power (to protect detectors), and the ability to open and close the telescope enclosure. A catastrophe (earthquake or major storm) might shut down mountain power for weeks, during which time people who might be stranded in the support building should have electrical power and other amenities. Tytler and Nelson (2002) recommend that 170 kW of emergency power be available.

Elevators
One or two elevators must be built to connect the enclosure floor to the Nasmyth platform. Each should be large enough to carry small items of equipment on carts. Each should have dimensions of 1.5-m door width, 2.1-m height, and 3-depth, and have at least a 3-ton capacity. These elevators might be fixed to the ground or attached to the moving structure. There are a number of advantages to an elevator attached to the telescope structure, but there are also consequent safety issues that must be considered. The choice of elevator geometry will depend on detailed analysis of the requirements and the geometry of the telescope and enclosure design.

Ventilation and Refrigeration
The enclosure shall provide an environment to the telescope that preserves the natural seeing of the site. This will involve heating, cooling, insulating, and ventilation. Some of these issues are discussed in Section 12.4. At night, ventilation is required to remove heat from the telescope to maintain the enclosure at the ambient outside air temperature. During the day, the closed enclosure must be cooled to the anticipated nighttime temperature. Heat sources that define the cooling requirement include sunlight conduction through enclosure surfaces, warm air infiltration, artificial light, instruments, telescope bearing oil, and telescope structure. Tytler and Nelson (2002) estimate the required refrigeration is about 400 kW of heat transfer.

Forced air ventilation of the enclosure, as is employed at Keck, may not be practical for an enclosure of this volume, due to the high power needs of the fans. We will explore this in more detail in the next phase of work.

Doors
A 6-m wide by 5-m high door is required for the entry of loaded trucks or a mobile crane.

Cranes
Cranes will be needed at the observatory during construction and operations. For flexibility we suggest that a large mobile crane will be sufficient for the following activities:
- reaching all parts of the inside of the enclosure
- reaching all parts of the outside of the enclosure and cylinder
- reaching all parts of the telescope structure
- lifting the heaviest moving parts of the telescope and enclosure, estimated at 90 metric tons

- installing and replacing secondary mirror
- installing and replacing clusters of primary mirror segments
- aiding in the construction of all buildings and facilities on the summit
- installing and servicing instruments on the Nasmyth platform

If the mobile crane can be used for all activities inside the enclosure (such as installing and replacing clusters of primary mirror segments), then we may not require any cranes attached to the enclosure.

12.2.2 Summit Building Requirements

The summit building must provide:

- facilities to support work on the telescope and its optics including construction, cleaning, maintenance, storage, engineering tests, and upgrades
- laboratories and storage space for the assembly, calibration and maintenance of astronomical instruments and AO systems
- a room for telescope control
- a room for AO laser equipment
- a room for communications and computing equipment
- personnel facilities (restrooms, kitchen, lounge, research/study/computing room, first aid station)

12.3 Reference Design Concept for Enclosure

There are many possible designs for the telescope enclosure, and these are reviewed Tytler and Nelson (2002). In order to reveal the complexities that determine the capital and operating costs we have adopted a reference design concept. In the preliminary design phase we will make a detailed comparison of alternative configurations and select a baseline design. The main features of the design concept follow.

- The enclosure comprises two parts: a *spherical metal dome* that rotates in azimuth, on top of a *fixed cylindrical wall*.
- The astronomical opening, through which the telescope looks, is set at the minimum acceptable size, a circle 16.25 m in radius. With practical shutters, the opening is probably a square 32.5 m on a side.
- The opening is closed with three shutters that move up and down. These shutters can present the minimal opening at all elevations 25°-85° and a slightly larger opening nearer the zenith. They can also open fully to present a wide-open slit for greater ventilation.
- Additional ventilation doors in the dome and base cylinder, actively controlled to match wind and temperature conditions. Currently we do not know the extent to which they will be needed.
- An elevator will connect the Nasmyth platform with the enclosure floor.
- A fixed tower/platform in the dome will provide access to the telescope top end.
- A large mobile crane will be used inside and outside the dome for access, construction, and moving mirrors and instruments.
- All functions that do not need to be housed in the enclosure will be in an adjacent support building.

Table 12-1 (Tytler and Nelson 2002) gives the dimensions of the reference concept enclosure.

Table 12-1. Nominal dimensions for the CELT enclosure

Parameter	Nominal Value	Assumptions
Shape of rotating dome	Sphere	Centered on elevation axis
Dome inner radius	42 m	Adequate for secondary
Dome outer radius	45 m	3 m thickness for strength
Height of elevation axis	28.0 m	Azimuth bearing 1 m above dome floor
Dome exterior height	73 m	
Minimum astronomical opening	16.25 m radius	Smallest size with field of view and room for laser launch telescopes
Angular size of minimum opening at 45 m	21.17°	Measured from center of dome
Range of elevation for unobstructed view of sky	25°-90°	Telescope requirements
Height of dome fixed cylinder	7 m	Clearance for fixed 5 m door
Outer radius of dome at cylinder height	39.8 m	Dome outer radius, cylinder height
Cylinder wall thickness	1 m	Adequate rigidity
Inner floor area	4,729 m^2	
Elevation angle of bottom of rotating dome	-27.82°	Measured to the 45 m wall from the dome center
Dome outer surface area	18,661 m^2	
Dome inner surface area	16,256 m^2	
Moving mass of dome (steel)	1567 tons	Scaled from Fukuoa steel dome
Moving mass of dome (aluminum)	909 tons	Temcor aluminum dome design
Dome interior volume	256,000 m^3	Measure to inner radius
Dome wall volume	58,850 m^3	
Total clear interior volume	289,000 m^3	Dome + cylinder

Most commonly, telescope enclosures are made of steel. At least two companies have experience making large fixed domes from aluminum space frames: Temcor (Carson, CA) and Star*Net (Longwood, FL). The Hobby-Ebberly Telescope dome is of this type. Temcor has presented a design for a basic CELT sized dome (Mixter and Porter, 2000).

Elevator

The elevator design concepts meets the requirements given in Section 12.2.1.

Tower

A tower is fixed to the north or south side (depending on site location) of the dome floor adjacent to cylindrical wall. Its purpose is to service the secondary mirror area. The top of the tower should be approximately 40 m above the dome floor, within 1 m of the bottom of the f/15 secondary mirror. Ladders and gantries will allow access to the entire front and back of this 4-m diameter mirror. This tower will also include an elevator similar to that reaching the Nasmyth platform.

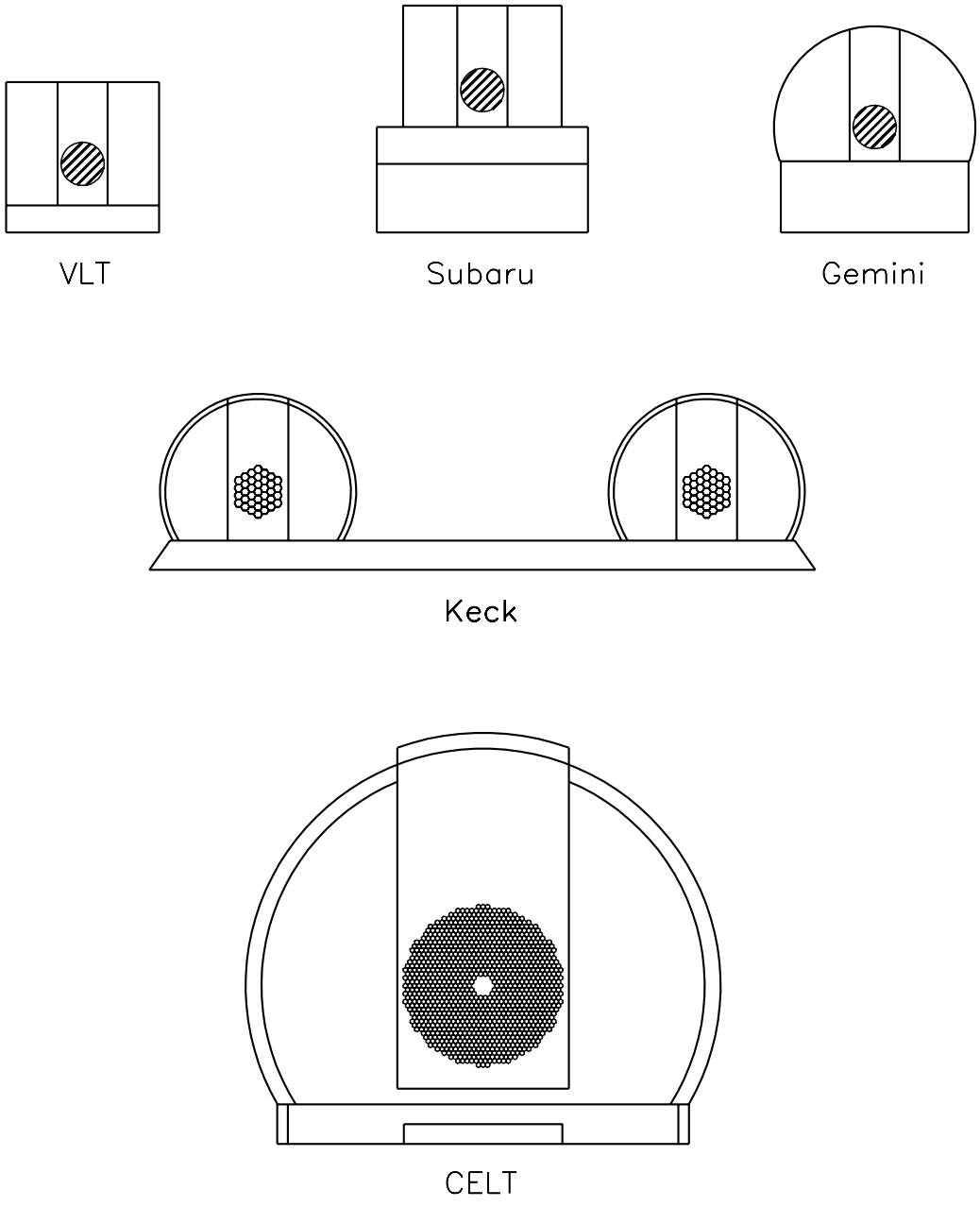


Figure 12-1. CELT enclosure concept and the enclosures for the Keck, Gemini, VLT, and Subaru telescopes drawn to the same scale. The primary mirrors are indicated.

Shutters
We have selected a three-shutter system to provide the minimal opening size and motion required. Each shutter moves independently up and down tracks on dome arch girders. They all have the same width and do not move over or under one another. Table 12-2 gives extension, open and closed heights, mass, and energy required to move from open to closed for each shutter. These values from Table 2 of Tytler and Nelson (2002) assume that the shutters have a lightweight space frame steel structure and a 3 mm thick titanium skin.

Table 12-2. Shutter concept parameters

Shutter size seen from dome center (degrees)		**Shutter center height range (m)**	**mass (ton)**	**Energy (MJ)**
top (C)	46.587	40-73	100	32
middle (B)	30.375	42-64	65	14
bottom (A)	30.375	18-43	65	16
			Total: 230	**52**

In order to achieve maximum wind shielding, at elevation angle $\simeq 54°$ the middle shutter must be moved across the telescope beam. Figure 12-2 shows the geometric features of the three-shutter design. Figure 12-3 shows the telescope and dome at 54° elevation angle.

We note that the Keck shutters have been a regular source of problems, and require high maintenance. We will pay special effort to ensure the CELT design is robust and will not be a maintenance issue. These efforts will include a careful study of the effects of structure deformations due to temperature changes, load variations, wind loads, etc.

12.4 Enclosure Thermal Issues

To not degrade the seeing we need to control the temperature of the enclosure actively and passively. Our goal is to keep the air in the light beam at the ambient nighttime temperature. The following heat sources are addressed in the enclosure design:

1. solar heat gain during the day
2. radiation heat loss during the day and night
3. convective heat gain during the day and night
4. heat gain from artificial lights
5. heat gain from telescope bearing oil
6. changes in nighttime temperature
7. latent heat load from moist air infiltration
8. instrument electronics heat load

We give some thermal features of the concept enclosure design and estimate the power impact of each of the above sources.

1. Solar heat gain during the day:
   We will use a reflecting paint on the outer skin to reflect solar radiation and will use insulation on the inside of the dome wall to keep the solar heat out of the dome interior. The insulation will reduce the transmission of the solar heat into the dome interior and limit the refrigeration needs.

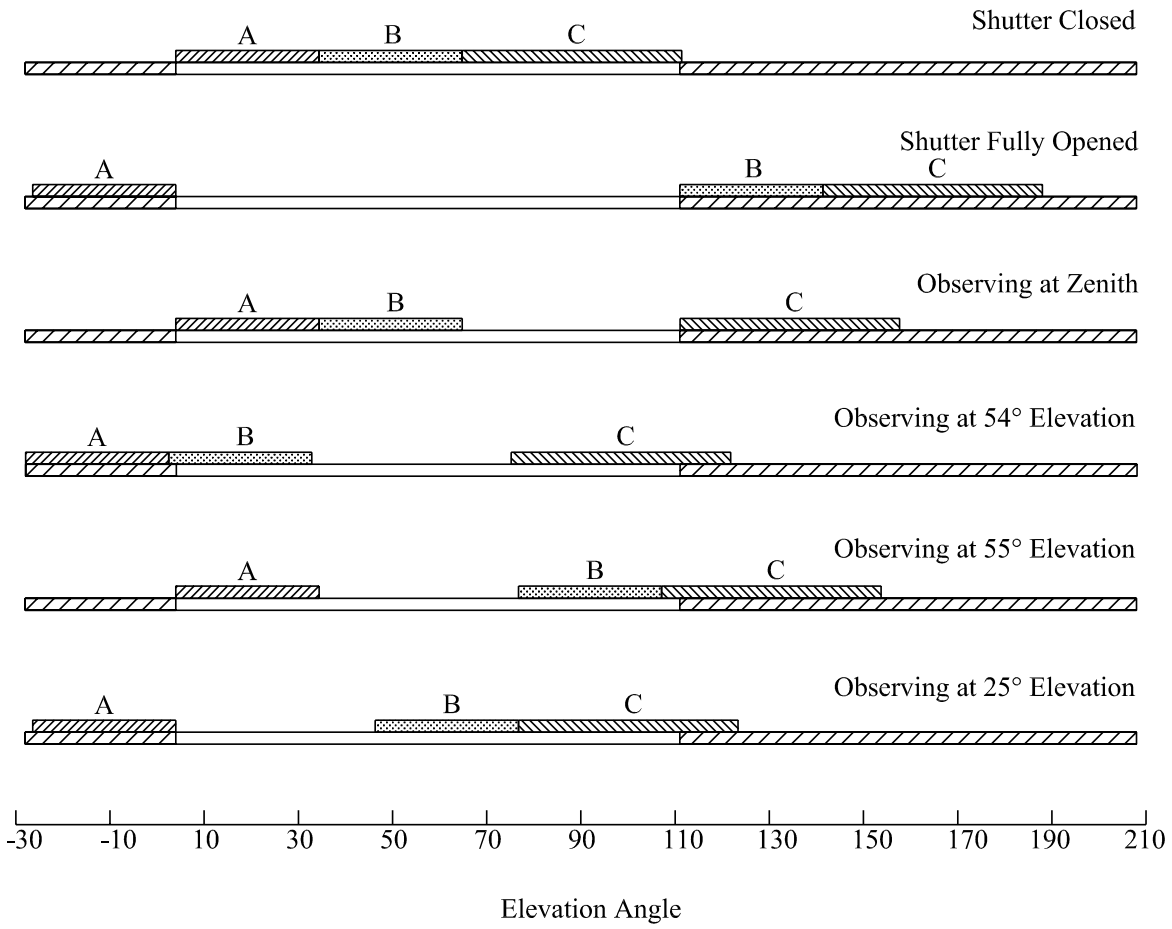


Figure 12-2. Representative shutter positions are shown for the three-shutter design. The dome outer surface has been flattened, so the horizontal scale is angle. The top drawing shows the dome and the three shutters in the closed position. The second drawing shows the shutters fully open. The third shows the shutters set for maximum wind protection while observing at the zenith. The fourth drawing shows the telescope at its minimum elevation angle before the middle shutter (B) must be moved up above the telescope. The fifth drawing shows the telescope at its maximum elevation with full wind protection and shutter B above the telescope. The final drawing shows the shutter configuration for observations at the minimum elevation angle allowed.

To estimate the magnitude of this source we assume:

- solar flux of 1000 W/m^2
- a reflectivity of 0.8, averaged over the solar spectrum (dirty white paint)
- air flow of 2 m/s which carries away 12 $W/m^2/C$
- ambient air temperature of 275° K
- area A = 0.4 of the total outer surface area (8,164 m^2)
- radiation to sky with effective sky temperature of 255°K

Under the above conditions, we find that the daytime dome skin temperature rises to 279°K. The nighttime skin temperature is about 272°K. We use Fourier's Conduction Law to calculate the power flowing through an insulated wall in steady state: $H = -kA\Delta T/L$. Assuming k = 0.04 W/m/K, the thermal conductivity of fiberglass, and a thickness L = 0.10 m gives H=13kW.

2. Radiation heat loss during the day and night:
   The dome and telescope will over-cool at night by radiation to the sky and convection to the ambient air. We will use low emissivity coatings on parts of the enclosure and telescope above the primary to minimize this cooling for surfaces above the primary. Power radiated through the slit will be dominated by the lower part of the dome and the floor. We will use aluminized Mylar reflective coating on parts of the dome near the slit to minimize cooling near the light path. If we assume the Stefan-Boltzman law, cavity emissivity of 1, dome interior of 275° K, the sky of 255° K, and a minimum square astronomical opening of 1056 m^2, then we calculate 89 kW radiated. This is an upper bound since we have ignored the effect of the low emissivity primary. Since the primary is deep in the dome, it is probably a 10-20% effect.

3. Convective heat gain during the day and night:
   Warm outside air infiltrating through seals is a major source of heat. We assume wind speed = 5m/s, seal gap width = 1 mm, outside temperature minus inside temperature = 5° K, air heat capacity of 1kJ/kg, and density of 0.76 kg/m^2, to calculate a heat infiltration of 19 W/m of seal. We assume that a 1 mm gap can be achieved with seals, which are inflated when parts are stationary. The infiltration at Keck is approximately 260 W/m, corresponding to approximately 14 mm gaps, which is unacceptable for CELT. Tytler and Nelson (2002) assume 81 square ventilation doors, 4.4 m on a side (these dominate the total seal length) to calculate a total seal length of 2100 meters. The total infiltration is then 40 kW.

4. Artificial lights:
   Tytler and Nelson (2002) estimate a maximum of 69 kW from artificial lighting.

5. Telescope bearing oil:
   We will precool the hydraulic bearing oil so that it does not add heat to the dome air. Tytler and Nelson (2002) estimate that about 100 kW cooling will be required.

6. Changes in nighttime temperature:
   If the predicted nighttime temperature changes (ΔT), then we will have about 12 hours to change the telescope to the new predicted temperature. Assuming the telescope mass = 2M kg, specific heat of steel = 470J/kg, and ΔT = 4 K; then we need 87 kW continuously for 12 hours.

7. Latent heat load from moist air infiltration:
   We must ensure that the dome air is dry enough to not condense onto the cold optics (at expected nighttime temperature). This dehumidification may require ~ 40 kW of power.

8. Instrument electronics heat load:
   Electronics from scientific instruments and the AO system may generate significant power and must be included in overall assessments of power, heat loads, and cooling requirements. These levels could be of order 100 kW. As the instruments become better defined we will assess their power and heat load implications.

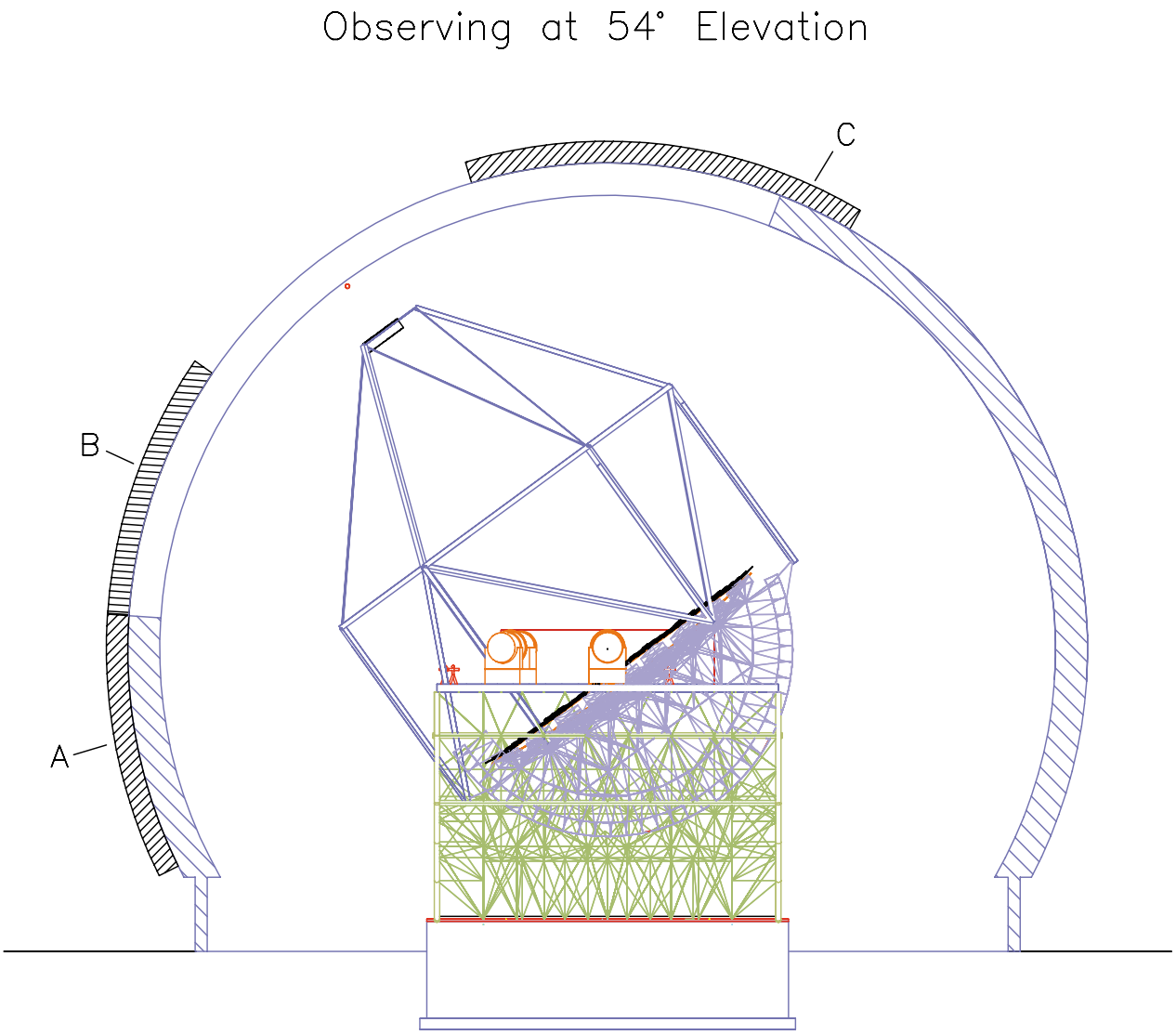


Figure 12-3. The telescope observing at 54° elevation is shown. Shutters A and B are below the telescope. Observing at lower elevation angles will require moving shutter B above the telescope.

12.5 Local Headquarters Facility Requirements

The local headquarters facility needs to support about one hundred personnel working to maintain and operate the observatory. The operations are discussed in CELT Report No. 35 (Nelson and Mast, eds., 2002). The local headquarters will require:

- facilities and laboratory to store and assemble primary mirror segments and clusters (possibly some of this could be at the summit) and other major optical subsystems
- a facility to maintain a fleet of vehicles for transport to and from the summit
- laboratories to make repairs when they are less expensive or more effective at the local headquarters

instead of the summit (this might include major repairs to scientific instruments and AO systems and the maintenance and repair of small telescope components)
- housing for visiting scientists
- facilities for maintaining active public relations with local communities
- computing services with appropriate power and air conditioning

12.6 Remote Observing Facilities Requirements

Remote observing facilities will be located at the major participating astronomy departments in California. Each will require facilities for nighttime observing that will include personnel support, telephone and video conferencing, and rapid data exchange. Personnel support for two to four people will be required. This will include quiet daytime sleeping accommodations for two and nighttime lunch facilities for four. Given the strong dependence on communications technology and the rapid evolution of that technology, we have decided to postpone the definition of specific requirements until a later stage of the observatory design.

12.7 Future Design Effort

Tytler and Nelson (2002) present many aspects of studies required for an integrated facility. It gives specific questions to be answered and a list of resources available to help answer them. We summarize here the systems integration studies and refer the reader to the report for the detailed questions and resources.

The report recommends using a combination of studies to determine the likely impact of the enclosure on the performance of the telescope:
- studies that were performed to support the design of existing enclosures
- studies of measurements already obtained at existing enclosures
- new measurements at existing enclosures, especially those similar to the enclosure that we envision for CELT (e.g., Keck)
- computer simulations of the telescope and enclosure including aero (wind), thermal (temperature), and elastic (vibrations, deformations) simulations with sufficient resolution to help design the enclosure
- wind tunnel tests on models of the telescope and enclosure

A combination of all these studies is recommended because the enclosure will have a major impact on the telescope performance, and will be a major portion of the total capital and operating costs.

These studies will address the telescope, enclosure, and environment (including wind, dust, humidity, vibration, and thermal) as a system in the following ways:
- We will calculate the performance of the telescope as it will be affected by the expected range of environmental conditions in any given enclosure.
- We will need tools to identify those changes to the telescope and enclosure that will improve the performance of the telescope.
- We will quantify the improvement in performance that comes from a given change in the telescope and enclosure, for example, how we should set the clearance between the telescope swing and the inner radius of a spherical dome.

- We will quantify the changes in capital and operating costs that come from changes in the telescope and enclosure.

In the next phase of design we will create a detailed requirements document. We will carry out extensive discussions with candidate manufacturers, support design studies, and select a reference design configuration.

REFERENCES

Mixter, T.J. (Temcor Corporation), D.S. Porter (National Optical Astronomy Observatories). 2000. "Next generation telescope dome enclosure a 91 meter (300') geodesic dome." SPIE Proceedings **4004,** 4004-54.

Nelson, J. and T. Mast, eds. 2002. "Operations Plan." CELT Report No. 35.

Tytler, D. and J. Nelson. 2002. "Buildings, Telescope Enclosure, and Facilities." CELT Report No. 23.

Chapter 13. Site Selection

13.1 Introduction

Site selection for CELT is a critical issue on many levels, and is already on the critical path.

Scientific productivity of a telescope is directly proportional to the number of good nights available (good in terms of the transparency, seeing, usability of the adaptive optics, etc.). The quality and reach of frontier science depends strongly on the typical and optimal conditions available, especially on seeing. For example, in terms of the limiting depth, one can to first order trade seeing diameter with the telescope diameter. Clearly, the world's best telescope should be situated on the world's best available site. To do otherwise would mean a waste of scientific opportunity and poorly spent cost and effort.

In addition, site selection has tangible consequences beyond its direct impact on science. It will strongly affect the cost and ease of the telescope construction and operation. It will affect the activities of management, technical support, and personnel recruiting. Some major equipment failures would lead to a downtime whose duration could depend on access and location of the telescope.

On a more subtle level, a detailed characterization of the site can affect the telescope and dome design (e.g., due to the wind speed distribution, mechanical properties of the soil, seismicity, etc.), and the adaptive optics (AO) design (through the various atmospheric turbulence properties). It is imperative that an intensive and immediate effort be devoted to the selection and characterization of one or more candidate sites.

As we will argue, it may be impossible to guarantee an optimal selection of a site because of the intrinsic atmospheric and climate variations that exist on all time scales. However, a reasonable selection can be made. The confidence in a particular site decision will be directly proportional to the quality and the time baseline of the testing data. With the anticipated short timeline for the CELT site selection process, this may be a fundamental problem. Adequate, homogeneous site comparison data (especially with AO issues in mind) simply do not yet exist for most potential sites. Fortunately, the CELT effort may be able to benefit from previous and ongoing work along these lines by the groups at National Optical Astronomy Observatory (NOAO), European Southern Observatory (ESO), and elsewhere. We can join with them as a part of a collaborative site evaluation program.

13.2 General Issues

Perhaps the key problem is the lack of comprehensive, homogeneous site testing data over long time baselines (years to decades). Heterogeneity (and non-simultaneity) of measurements makes a fair comparison of different candidate sites difficult. Measurements over a short time baseline may lead to the selection of a site that later experiences poor conditions for years and decades. Perhaps this bias is statistically unavoidable, but it should be recognized. Most of the site testing done more than a decade ago did not address key issues relevant for the design and operation of AO systems. To some extent the same is true of IR site qualities. These are precisely the two directions of observing technology and practice we now anticipate will be most important for large telescopes and astronomy in the years ahead.

Finally, it is not clear that we have today a complete list of candidate sites to be tested and compared. In addition, most site selection efforts in the past were limited in scope and focus, e.g., with visual estimates of the seeing and cloud cover, and were based on very short testing campaigns. They also included considerations that we now consider to be less important or even undesirable, such as the proximity to

a major urban center or university within the U.S. It is likely that the only sites for which even remotely usable, relevant, modern data exist are those where there are already major observatories. At the end of this chapter we list some pertinent literature and Web sites.

Conventional wisdom is that there are two types of superb sites in terms of atmospheric properties:

1. Isolated high mountains on islands in temperate oceans, where the weather is good and a laminar air flow and large thermal inertia of surrounding ocean keep the inversion layer low. These guarantee good seeing. There are two locations known on this planet: Mauna Kea in Hawaii and La Palma in the Canary Islands. See Walker (1971) for a list of the other possibilities.

2. Coastal mountain ranges next to a cold ocean current with stable subtropical anticyclone conditions. The proximity to the coast allows for unperturbed laminar air flow. The cold ocean, whose influence may extend to some distance inland, keeps the inversion layer low. These conditions exist on the coasts of California (including Baja), northern Chile, and Namibia.

While it is generally believed that continental sites would be inferior to the above, mainly due to turbulence-generating topography, the situation is not clear. An example demonstrating that superb sites (at least in terms of seeing) do not all fit this pattern is Maidanak in Uzbekistan (Ehgamberdiev, et al., 2000). It is possible that some superb and so far undocumented sites do exist in, for example, the U.S. southwest, northern Mexico, or the high mountains of northern Africa.

13.2.1 Astronomical and Technical Issues

The most basic issue determining the usefulness of a given site is the fraction of nights during which observations can be made. For a superb facility like CELT, we should be more restrictive, and consider only the number of high-quality nights (clear, good seeing), since those are the nights when the telescope can be pushed to its limits and cutting-edge science can be done. Backup observing programs that can be carried out in mediocre or poor seeing and/or non-photometric conditions are not the science drivers for CELT. We also note that AO using laser guide stars, which may be the key operating mode of CELT, requires at least reasonably photometric conditions.

Given a reasonably high fraction of high-quality nights, probably the single most important issue for ground-based UVOIR astronomy is the quality of the seeing. This depends on the overall atmospheric turbulence properties of the site. Seeing quality also affects the design and performance of the AO systems. A simple way of characterizing the seeing is through a distribution of measured image sizes (FWHM). A comprehensive way to characterize turbulence is with a distribution of C_n^2 profiles, from which all other characteristic parameters can be derived (see Section 13.3.2).

The next important issue is sky brightness and transparency in the mid-IR, determined by the percipitable water vapor (PWV). This is the key factor in selecting sites for mid-IR to mm wavelength telescopes. We note here that the only atmospheric turbulence that matters at such wavelengths is the turbulence that occurs in the water vapor layers. The dry air turbulence is irrelevant. However, dry air turbulence is important for observations at UV to NIR wavelengths. Thus a site which may be superb for mid-IR to mm astronomy may not be as good for OIR astronomy, the focus of CELT. Thus, we need to make a scientifically-motivated choice between a site that provides a sizeable fraction of mid-IR-capable nights and a site with a superior total fraction of clear nights or mean seeing (but with a high mean PWV).

In addition to its role in seeing and AO-related issues, the distribution of wind speeds is also important for telescope and dome design. It is possible that some otherwise attractive sites may be too windy to effectively operate a telescope of this size.

It is important to study the climate issues. Strong fluctuations in weather patterns and phenomena are seen on all time scales for which data exist, spanning the few decades where modern meteorological data has been collected, to a few centuries where historical records exist (tree rings, "mini ice ages," etc.), and out to millions of years where Greenland and Antarctic ice cores provide the record. There is no way to guarantee that any site selected objectively on the basis of only a few years of testing data will remain good a decade or more.

Global warming is an example of a secular trend. While this topic is surrounded by political controversy, there is a broad scientific consensus on some general features. Two features that may be of interest in the present context (a time scale of a few decades, the scientific lifetime of the observatory) are:

1. A gradual expansion of the climate zones from the equator towards the poles, including wet tropical zones and dry subtropical desert zones in the Americas. This may be a factor in selecting a site in Chile, the U.S. Southwest, or Mexico. Hawaii weather will probably be relatively unaffected.

2. A perceptible increase in the frequency and strength of extreme weather events, e.g., major storms. This will affect all of the areas under consideration, including Hawaii. Extreme weather events impact the structural design requirements of the telescope, the dome, and other summit support buildings; and they will result in the loss of observing time.

Another issue of scientific importance is available sky coverage. Sites at lower geographic latitudes are better in this regard. However, this is only critical for all-sky surveys, which CELT would not be doing, and for coverage of very rare and interesting types of objects or phenomena (for example, a supernova in the Magellanic Clouds, or in M31). We conclude that this issue should not be a strong driver in site selection.

An additional concern is light pollution. This has primarily been important in traditional seeing-limited, visible light applications. This problem is greatly diminished in the IR and with AO, precisely the directions in which CELT astronomy is expected to move. (One could even argue that if we were to abandon the seeing-limited visible regime, then Mt. Wilson may be a superb candidate site.) However, most sites under consideration are remote and as a consequence are expected to be relatively free of this problem. A comprehensive study was published by Cinzano, et al., (2001a, b).

Finally, geological issues also play a role. The seismic environment is important. The telescope, instruments, AO systems, enclosure, and building designs must all be earthquake-safe and robust. The vibration environment during operations will strongly impact the telescope, AO systems, and instrument optics control. Nearby active volcanoes may present a physical danger and cause a serious loss of observing time due to ash fall and extinction following eruptions. These are low probability, but high impact, problems. Unfortunately, nearly all sites and areas under consideration are located in geologically very active areas, so this is a factor to consider.

13.2.2 Logistical Issues

The geographical location and local topography of the site are critically important for the logistics of construction and operation. They could possibly affect the overall cost by as much as a factor of two for the extremely remote locations.

These factors will also impact physical access and travel for people and heavy equipment, and the availability of supplies and resources. A key concern is the existence of and proximity to the infrastructure grids of roads, utilities (water, power), communication networks, etc. These strongly depend on whether the telescope is built on an already developed site or on a newly established site. In the latter case, which almost certainly applies for remote locations in Chile and possibly also in Mexico, the cost of providing access roads capable of moving heavy equipment may be prohibitively high. In any case, we will need to provide all such infrastructure local to the telescope. In the case of building and operating on an already developed site, the existing infrastructure will almost surely have to be significantly upgraded.

The site altitude impacts the costs of construction and operations. Work will be harder, less efficient, and more costly at the higher elevations (e.g., in the Chilean Altiplano, including Chajnantor, versus the coastal mountain ranges). For altitudes commensurate with Mauna Kea or higher, an intermediate-level base camp, analogous to Hale Pohaku, would have to be developed.

Another key issue is the proximity to adequate medical care facilities, especially for the remote, high-altitude sites, since accidents are statistically certain to happen. The history at Mauna Kea shows that facilities must be prepared to address severe cases of altitude sickness. The CELT observatory will require these facilities. They may need to be included as part of the observatory construction.

Site location will also affect the availability of qualified technical personnel for both construction and operation phases. For an advanced facility such as CELT, it is highly likely that a considerable number of highly qualified technical personnel will need to be present at the site, especially in the early years of operation. A more remote and/or high-altitude site will almost surely require a tour-of-duty system, even with its attendant inefficiencies. The site will determine the management model, including the number and the location of facilities, etc. Finally, for a technologically complex instrument like CELT, a reliable and rapid supply of spare parts and similar items will be essential.

13.2.3 Political Issues

Political climate and stability are also important issues, both in the short term (permission to prepare site and construct) and long term (operations over an anticipated useful scientific lifetime of several decades). Unfortunately, political and meteorological climates seem equally unpredictable, both in the U.S. and elsewhere.

The concerns range from basic administrative issues to the physical safety of the observatory and personnel. These concerns arise from the goals and actions of governments (for the sites outside the U.S.), ethnic groups, and leaders of political and/or social movements (nationalistic, environmental, religious, etc.). The delicate situation in Hawaii is well known. The situation in Chile appears to be reasonably stable and friendly, as long as relations with the government and local community are handled with care and respect .

Regardless of the choice, it is paramount that the local community, including the astronomical community (especially for site choices in Chile or Mexico), is involved in the project from the early stages. This involvement must be friendly and constructive. The community must be both partner and participant in the project, and must derive pride and many other long-term benefits from its participation. To these ends, it will be necessary to develop educational and local employment programs. These will require project management time, financial support, and possibly observing time.

Finally, the site location must be acceptable to the relevant funding agencies (including the State of California) and potential donors. Precedents exist to build such a telescope outside of California (e.g., in Hawaii) and outside the U.S. (e.g., in Chile); but an assurance should be obtained from the donors and relevant funding agencies prior to final site decision.

13.3 Site Characterization and Testing Methods

As we emphasized above, the comprehensive and homogeneous data needed to make an optimal site choice simply do not presently exist and realistically cannot be obtained in the short time available.

After one or more candidate sites are selected using climatological and geographical criteria, then detailed numerical modeling (of the type done by De Young and Charles, 1995) can be used to supplement direct measurements of the atmospheric parameters at the site. However, there is no adequate substitute for actual onsite measurements spanning at least one full annual cycle.

13.3.1 Weather

Standard meteorological data, including distributions and diurnal changes in ambient temperature, humidity, wind speed and direction, precipitation, cloud cover, etc., are routinely collected at all existing observatories, and even at some popular, as-yet undeveloped candidate sites (e.g., Chajnantor area in Chile). Portable meteo-stations are easily obtained and can be deployed to other sites to be tested.

A powerful methodology for uniform and objective selection and comparison of candidate sites over wide areas is the use of remote sensing (satellite) data, which can cover a period of about 10 years with a few-hour sampling rate. An extremely useful study of this type was already completed by Erasmus and van Staden (2001) for the site selection process in Chile, sponsored by NOAO and other partners. A restricted study of several sites in the U.S. Southwest was also done (Erasmus 2000). We anticipate its expansion to cover all of the relevant area in the North America and to include a comparison study of Mauna Kea. These studies may be the best and only available homogeneous comparisons of a large number of candidate sites. Unfortunately, these studies are limited to two parameters, the fraction of clear (and partly clear) nights, and the PWV fraction. In addition, they have a limited spatial resolution. Their primary use is to compile an objective list of candidate sites, leading to more detailed in situ measurements of other parameters.

It is also possible to estimate probable seeing quality from meteorological data (at least on a coarse grid) that can be used to rank individual mountains, but not specific sites on a given mountain. This can be done through analysis of atmospheric flow structure. In the free atmosphere, wind speed and direction changes with height create turbulent layers that degrade seeing quality. From SCIDAR measurements it is known that turbulent layers are formed near the jet stream level (at ~ 200 mb, i.e., ~ 12 km altitude) and at the boundary of the upper westerlies and the low-level circulation (e.g., winds, topographically produced drainage flows, land-sea breeze, etc.). The strength of the upper turbulent layer is related to

wind speed shear above and below the jet stream. For the lower layer it is related to the magnitude of speed and direction changes across the boundary. Additionally, a strong vertical temperature gradient is typically observed coincident with these wind shear layers. Relationships between turbulence parameters (e.g., Richardson number, C_T^2, and C_n^2) and vertical gradients of wind and temperature have been defined empirically and theoretically. Knowing the vertical structure of the wind and temperature field over an area where potential telescope sites exist, it is possible to map the frequency of occurrence, relative strength and height of turbulent layers that degrade seeing.

13.3.2 Astronomical Seeing and Atmospheric Turbulence

A general description of the problems caused by the atmosphere is given in the Adaptive Optics Chapter 9. The measurement issues are described in more detail in the report by Schoeck (2001), along with a list of references. Here we summarize some of the key points. Tables 13-1 and 13-2 summarize the different instruments and methods, and their scope, cost, and requirements.

Table 13-1: Instruments and observed objects required for different atmospheric characterization methods

Method	Instrument	Object
SCIDAR*	1 large aperture (≥ 1m)	double star
generalized SCIDAR	1 large aperture (≥ 1m)	double star
MASS**	one 20 cm subdivided aperture	single star
Scanning scintillometer	small aperture, single-element detect	single star
Scintillometers	small apertures (10 – 40 cm)	diverse
DIMM***	2 small apertures usually part of larger ap. of D ≤ 50 cm 1 CCD	single star
GSM****	4 small apertures (10 cm) 4 APDs	single star
WFSs*****	WFS (usually part of AO system)	single or double stars
Interferometers	existing interferometer	single or double stars
PSF analysis	large telescope	single or multiple stars
Balloons	micro-thermometers, velocimeters	none
Acoustic Probe	Acoustic sounder	none

* Scintillation detection and reading
** Multi-aperture scintillation sensor
*** Differential image motion monitor
**** Generalized seeing monitor
***** Wavefront sensors

Table 13-2: Measurable quantities and sensitivity to ground layer turbulence

	$C_n^2(h)$	v(h)	resolution	<v>	r_0	τ_0	L_0	θ_0	ground layer
SCIDAR	yes	yes	low	yes	yes	yes	no	yes	no
gen. SCIDAR	yes	yes	low	yes	yes	yes	no	yes	yes
MASS	yes	no	low	maybe	yes	yes	no	yes	yes
Scanning scint.	yes	maybe	low	maybe	yes	yes	no	yes	yes
Scintillometers	yes[a]	yes[a]	low	yes[a]	yes[a]	yes[a]	no	yes[a]	(yes)[a]
DIMM	no	no	—	no	yes	yes	no	no	yes
GSM	no	no	—	yes	yes	yes	yes	yes	yes
WFS	maybe[a]	maybe[a]	low	yes[a]	yes[a]	yes[a]	maybe[a]	yes[a]	yes[a]
Interferometers	no	no	—	no	yes	yes	yes	yes	yes
PSF analysis	no	no	—	no	yes	no	no	yes	yes
Balloons	yes	yes	high	yes	yes	yes	yes	yes	yes
Acoustic sounder	yes[b]	yes[b]	high	-	-	-	-	-	yes[b]

[a] Not all of these quantities can be measured with all kinds of scintillometers or wavefront sensors.
[b] The acoustic sounder typically probes only the lower atmosphere, up to 1 km.

A quantitative description of turbulence for site evaluation purposes only makes sense in statistical terms. However, as atmospheric turbulence varies on almost all temporal and spatial scales, it is not only important to know the mean values, but also the deviations (variances, probabilities, time and spatial scales, etc.) from the mean.

The statistical properties of turbulence are described (more or less completely for our purposes) by the C_n^2 profile, the wind velocity profile, and the power spectrum of the wavefront phase (or a related quantity like angle of arrival or refractive index). The C_n^2 and wind profiles need to be measured for all sites of interest. Simply determining an overall parameter like r_0 is not sufficient for extremely large telescopes (ELTs). The power spectrum will probably need to be measured in order to determine the deviation from the Kolmogorov spectrum for large spatial scales. These deviations are likely to be important for an ELT like CELT, and it is not a priori clear which model is appropriate to describe this large-scale behavior.

We list below the quantities of interest. It might not be necessary to measure all of these quantities separately since they are not all independent.

r_0: Fried's parameter, can be measured with any of the methods listed in Table 13-2.

L_0: The outer scale of turbulence can be measured with some methods, but interpretation is not trivial. It is not a priori clear which model should be used for the power spectrum.

C_n^2 profile: This is only measured by some of the techniques. Vertical resolution can be a problem, and it is currently unknown what resolution is required.

θ_0: Isoplanatic angle.

Wind speed profile: This is only directly accessible with balloons, but can be inferred using other methods and the assumption of the frozen flow hypothesis. The latter is not always valid for the spatial and temporal scales with which we are dealing.

τ_0: The atmospheric coherence time can be measured directly with a fast-frame-rate method, or it can be calculated from the turbulence profile, the wind speed profile, and the assumption of frozen turbulence.

Properties of the Na D layer: These can be measured by monitoring of laser star images.

There are many methods that can be used to measure the above quantities, including:

Scintillometers: There exist many methods of atmospheric turbulence characterization that are based on the scintillation of observed light. Some of the currently popular are SCIDAR, MASS, and the scanning scintillometer.

The value of r_0 can be determined from the variance of the magnitude of the scintillations. If one wants to determine turbulence profile, one has to use double stars, spatial filters permitting the measurement of different spatial frequencies, multiple or sub-divided detectors, or a combination thereof. In principle, both C_n^2 and velocity profile can be found, and thus almost all parameters of interest are accessible, although not necessarily with all types of scintillometers. However, these instruments are not sensitive to ground-layer turbulence unless the detector plane is conjugate to an altitude significantly different from the instrument altitude.

The SCIDAR technique analyzes spatial and/or temporal autocorrelations of short-exposure images of the scintillation pattern produced by a double star. The detector is usually a CCD behind a large-aperture telescope. The advantage of this method is that the vertical turbulence profile is accessible. Resolution is on the order of hundreds of meters, depending on binary star separation, wavelength, and altitude probed. Disadvantages include the need for a large aperture (> 1 meter), and the cost of construction and operation (both of which can be solved through a collaborative effort).

Tokovinin (1998) has proposed multi-aperture scintillation sensor (MASS). It is based on scintillation measurements using a small telescope (approximately 20-cm diameter) with the aperture being divided into a small circular aperture surrounded by 3 annular apertures. The signal is the correlation of the ratio of short-exposure intensities in the apertures. MASS is insensitive to ground layer turbulence. It is a simple instrument using a small aperture, but it has a low vertical resolution.

Wavefront-sensing techniques: Some instruments measure the wavefront phase while others measure the wavefront angle-of-arrival. The two most promising are differential image motion monitor (DIMM) and generalized seeing monitor (GSM).

DIMM was developed by Sarazin and Roddier (1990) for the very large telescope (VLT) site evaluation campaign. A standard DIMM uses a small telescope (~ 0.5 m aperture) with a two-aperture mask and some kind of optical element (beam splitter, prism, etc.). This produces two well-separated images, one for each subaperture, of the same single star on a detector (usually a CCD). The statistics of relative motion of the two sub-images is directly related to r_0. It is a simple, low-cost instrument, easy to automate, and not sensitive to tracking errors, vibrations, etc. The DIMM is now widely used, but it only measures r_0.

GSM was developed as an instrument to measure the outer scale of turbulence, L_0. It was later improved to measure r_0, τ_0, and θ_0. The GSM consists of 4 small telescopes (D ~ 10 cm) arranged in a cross-shape (maximum baseline ~ 1 m) all pointed at the same star. Each telescope uses a photomultiplier in combination with a grating to measure the angle-of-arrival of the incoming wavefront at approximately 200 Hz (see, e.g., Ziad, et al., 2000). It is a relatively simple and easy-to-use instrument. It could be made to operate automatically, and it can measure a large number of parameters. However, it does not provide turbulence profiles, and L_0 is measured using a short baseline (~ 1 m), so the results are model-dependent. Currently only one GSM device exists.

In general, most scintillation methods can be implemented as wavefront sensing techniques and vice versa. One simply has to exchange a scintillation measurement by an equivalent wavefront measurement. Turbulence at all altitudes (including the ground layer) can be measured with the same precision. If one has a full-scale wavefront sensor (WFS) available (for example at an adaptive optics system), measurements with both high spatial and high temporal resolution can often be taken. However, the wavefront is generally much more difficult to measure than scintillations, making necessary a more complicated and expensive instrument.

Interferometers: The great potential of using interferometers is that their baselines are comparable to or exceed the outer scale of turbulence. Outer scale measurements using interferometers should therefore be more reliable than measurements obtained with smaller instruments. The problem with this method is the need for a working astronomical long-baseline interferometer. Examples include the Palomar Testbed Interferometer (PTI) and the Keck Interferometer.

Balloons: Balloons can carry micro-thermometers and anemometers to measure the C_n^2 and wind profiles, and thus virtually all atmospheric parameters of interest. The altitude resolution is excellent, of the order of 10 meters. A balloon needs two to 4 four hours to go through the entire atmosphere. Therefore, the profiles measured are not instantaneous. Only a few profiles can be taken.

Acoustic and Radar Soundings (SODAR and RADAR): Sound reflection can measure density variations in the atmosphere and thus measure the temperature structure constant profile, $C_T^2(h)$, from which $C_n^2(h)$ can be inferred. The problem is that this technique only works up to altitudes of generally less than 1 km, making it of limited interest for astronomical purposes. The radar reflectivity from the atmosphere is also related to C_n^2 and can be used to measure the C_n^2 profile. The advantage of radar remote sensing over acoustic soundings is the larger range of 10 - 20 km, which is sufficient for astronomical purposes. However, the spatial resolution associated with these ranges tends to be poorer than that provided by the acoustic sounders. Radar measurements measure the radar refractive index, which is influenced by temperature and humidity effects. The magnitude of the humidity contribution must be known to predict effects at optical and infrared wavelengths.

Point Spread Functions (PSF) and Speckle Interferometry: The PSF width of long-exposure images is a direct measure of seeing and thus of r_0. If several point sources are present in the field, PSFs can be used to measure anisoplanatism. Short-exposure images can be also be used. It should also be possible to do a temporal analysis from short exposure images of point sources (using speckle development or "speckle boiling"). It is a simple method that produces lots of data, but it only measures the r_0 and τ_0.

13.3.3 Infrared Observing Considerations

IR sky brightness and transparency depend strongly on the PWV, and the dependencies increase with the wavelength, becoming critical in the mid-IR. For this reason, telescopes covering mid-IR to mm wavelengths tend to be located at high and dry sites. An example is the Chilean Altiplano area, including Chajnantor, where ALMA and the proposed Cornell University telescope would be located. Otherwise, the considerations are identical to those for the visible light regime (seeing, fraction of the clear nights).

13.3.4 Optical Sky Brightness and Changes

Studies of sky brightness by Cinzano, et al., (2001a, b) and Garstang (1989a, b; 1991) are probably adequate for our present needs. Historical records of changes over a period of years to decades also exist for most of the well-established observatories. Visible and IR night sky afterglow brightness is also modulated by the Solar activity cycle.

13.4 Available Data for Some Candidate Sites

We do not want to suggest an overly pessimistic picture regarding the available site testing and comparison data. A considerable body of heterogeneous data does exist, although mainly for the sites with operating observatories.

A quick summary is that the most likely sites for CELT are either Mauna Kea, or one of the many possible sites in the northern Chile; but there may be viable alternatives in the southwest U.S. or northern Mexico. While other competitive sites may exist elsewhere in the world (e.g., in Namibia), they may be impractical for the logistical reasons outlined in Section 13.2.2.

In the following subsections, we avoid mention by name of any specific new candidate sites. We note that a number of viable possibilities do exist.

13.4.1 Mauna Kea

Mauna Kea (MK) is known to be one of the best astronomical sites in the world, mainly due to the superb seeing, at least on the summit and ridge. The high altitude also assures good IR observing conditions, and limited urban development assures a dark sky in the visible. However, in terms of the fraction of clear nights (both photometric and "spectroscopic", i.e., partly cloudy), MK is not exceptional and is almost surely inferior to many sites in Chile, and possibly elsewhere as well. It is also a very windy site, which may be a problem for a telescope as large as CELT. There are logistical benefits deriving from the fact that it is part of the U.S. and in an area attractive for living. Both of these make it easier to recruit expert staff. An extensive summary of the site properties of Mauna Kea can be found in Sarazin (2002).

In the deficit column there are political problems concerning the construction of additional telescopes on the mountain. While there is considerable uncertainty concerning the time scales of interest here, the Mauna Kea Master Plan (see the Web site below) describes the current vision and consensus. The Plan allows for construction of a single new ELT on the slopes N-NW of the summit area and the location of Keck Observatory. It does not appear that any scientific or technical considerations were involved in selecting this particular area.

Some information about this site does exist. It was tested by Walker (1983), who called it "North 1," and compared to a site very close to the present location of the Keck telescopes, which he called the "NW Cone." While the seeing measurements were done by eye by experienced observers, and the meaning of the units in which these measurements were expressed is not clear, the data are internally consistent and homogeneous. They include simultaneous measurements on the two sites, so that a good relative comparison can be made. A test comparison of the NW Cone with seeing estimates made at the Canada-France-Hawaii Telescope was also made. McInnes (1981) achieved consistent results.

The conclusion of this study is very clear: The North 1 site has seeing which is on average about 50% worse than the NW Cone. The differences are especially strong in the times of the good seeing (see his Table IX and his Figure 8). Walker attributed the seeing difference to the differences in ground cooling. The North 1 site is somewhat shielded from the wind relative to the summit and ridge cones, and would develop a ground turbulence layer. This is consistent with the observed seeing dependence on the wind strength and direction.

Walker further notes that there were a number of nights when the lower altitude North 1 site was in the clouds, while the summit and the ridge were above the cloud layer (a phenomenon familiar to many Keck observers). This suggests that an inversion layer may exist on some nights, below the summit level locations (e.g., the Kecks) but above the designated ELT site.

This finding was further confirmed by a study by Erasmus and Barnes (1989) (see their Figure A1). They made exhaustive echosonde measurements of C_T^2 profiles at the summit and two cinder cones. They also provide revised tower-microthermal measurements of C_T^2 profiles at the summit, Puu Poliahu, and at a site they call 13 North (just north of the designated ELT site in the Mauna Kea Master Plan). The tower measurements reach out to 30 meters above the ground. These data support the findings by Walker. Whereas the values of C_T^2 drop dramatically with the height for the summit/ridge and Puu Poliahu cinder cone sites, they stay nearly constant at the 13 North site, guaranteeing worse seeing.

If these measurements are confirmed and extended by a more modern, systematic comparison study of the designated ELT site with the existing telescope sites on the summit/ridge, this would argue strongly against placing CELT (or indeed any new telescope) at this location.

The Gemini/NOAO group is planning to perform simultaneous seeing measurements at different locations on Mauna Kea using two DIMM devices borrowed from Cerro Tololo Inter-American Observatory. This would also provide a data set that can be meaningfully compared with the measurements made in Chile using the same equipment. We anticipate that additional testing using other equipment would be needed. We recommend that CELT take an active role in making these measurements. A clear comparison of the designated ELT site with the summit sites (e.g., the Keck telescopes) is an urgent priority.

13.4.2 Chile

There is a (probably justified) perception among astronomers that good sites in Chile are superior to sites on Mauna Kea (and all known continental U.S. sites) in terms of photometric quality, but that Mauna Kea may be superior in terms of seeing. There is also a considerable range in the quality of sites in Chile, including the existing observatories. For example, ESO is gradually abandoning their old facility at La Silla (which is also probably comparable to Las Campanas, due to proximity) and has

developed at great expense the new location at Paranal. Neither of the sites used by NOAO and Gemini (Tololo and Pachon) seems to be as good.

Among the newly considered sites, Llano de Chajnantor, a plateau in the Atacama desert in northern Chile, will be the location of the Atacama Large Millimeter Array and is already a site of other high-frequency radio telescopes, such as Cosmic Background Imager. It is the anticipated site of the planned Cornell University mid-IR telescope. This general area has been designated as a National Science Preserve by the government of Chile, which greatly simplifies the site acquisition and approval issues. Results of a very useful early testing program for selected sites in Chajnantor have been published by Giovanelli, et al., (2001a, b).

While the site certainly seems very good for the mid-IR to mm wave observations, there are some troubling issues. For example, Giovanelli, et al., (2001a) find seeing distributions which differ by about 50% between two close sites that differ by only 100 meters in altitude (Chico and ALMA site). We note that the experience of Caltech's CBI group suggests that extended periods of bad weather can be a serious problem and may be amplified by global warming trends. This may be the so-called "Bolivian winter" phenomenon, where large masses of moist air from the Amazon basin are pushed over the Andes and deliver abundant snow on the astronomical facilities in the Altiplano .

Astroclimatology of La Silla (which is probably a good proxy for the Observatories of the Carnegie Institution in Washington site on Las Campanas nearby) and Paranal is described in detail on the ESO Web site (see below). While Paranal has undergone some deterioration since the start of the VLT construction, it is now recovering, and is almost certainly superior to La Silla. The mean seeing is not as good as on MK, but the fraction of the clear nights is higher. Neither Tololo nor Pachon seem to be as good as these sites.

It is likely that a new site would have to be developed for an ELT in Chile. NOAO and ESO are conducting a comprehensive program of candidate site selection and testing in Chile (see the Web sites below). This includes the study by Erasmus and van Staden (2001), and will be followed by detailed in situ testing of selected sites. To summarize some of the pertinent results from the Erasmus study, there is a suggestion of a bimodal distribution of the best sites. The best sites include:

1. Those with the highest fraction of clear nights and with a high PWV. These tend to be in the coastal mountain range, with typical altitudes of 2 - 2.5 km. An example is Paranal, the site of the ESO VLT.

2. Those with a low PWV and a significantly lower fraction of clear nights. An example is the high (typical altitudes 5 - 5.5 km) peaks in the Altiplano area (e.g., near Chajnantor). There is also a correlation between the typical wind speeds and altitude, and these high peaks may be too windy from the viewpoint of telescope mechanical design.

Adding to this dichotomy are differences in cost and ease of construction and operations.

This study by Esamus and van Staden (2001) is an excellent first step in selecting a short list of candidate sites in Chile. However, this type of study cannot measure the seeing or the quantities relevant to AO, which have to be measured from the ground. CELT participation in such testing of some of the best candidate sites found in this study, presumably in collaboration with the NOAO/Gemini and ESO groups, is a high priority.

13.4.3 North America

Most of the existing observatory sites in the U.S. were selected in the days before modern testing techniques and criteria even existed. Even those selected in the second half of the 20th century were based primarily on their suitability for seeing-limited, visible-light astronomy. None were selected for their suitability for the IR work or AO, although some turned out to be good for them. Other, non-scientific factors (such as the location in a given state, proximity to a particular university, etc.) influenced the choices. Urban growth has deteriorated the visible light night sky brightness for most of those sites. It is probably fair to say that none of the existing major observatory sites are as good as Mauna Kea or many sites in Chile. This, however, does not preclude the existence of superb sites that have not yet been studied. Maidanak in Uzbekistan is proof that superb sites do exist that do not conform to the conventional wisdom described in Section 13.2, and others may exist in North America.

Possibly the most relevant studies include those by Walker (1970, 1971), which led to the site selection for the Keck telescopes; and the studies undertaken in 1980's for the proposed National New Technology Telescope (Merrill, et al. 1986; and especially Lynds and Goad 1984). Walker identified Junipero Serra Peak as the most promising site among those he tested in California, but several other interesting possibilities exist. Mt. Graham in Arizona was favored by Steward Observatory and other partners, and is clearly a good site, but not superior to Mauna Kea. A number of other interesting candidates were suggested by Lynds and Goad (1984), but most of them lack site testing information. An interesting historical account of seeing measurements on Mt. Wilson was published by Teare, et al., (2000). Aside from the problem of scattered light from Los Angeles, there appears to be a secular deterioration trend in the mean seeing there.

A number of good sites may be found in northern Mexico. One example is San Pedro Martir (see Echevarria, et al., 1998, and references therein).

Satellite data studies, analogous to those done for Chile by Erasmus and van Staden (2001), of the possible sites in the North America can lead to an objective selection of a short list of candidate sites. These can then be tested for seeing quality and other parameters. The first step in this direction was the study by Erasmus (2000), and a much more extensive study is now under way.

It is also likely that the political and administrative considerations for the U.S. sites will be very complex.

13.4.4 Elsewhere

There are several other areas where superb (or at least promising) astronomical sites are known to exist, but for practical and logistical reasons we do not consider them among the primary choices for the CELT.

Canary Islands include sites of several observatories: For reviews of the testing programs, see, e.g., McInnes and Walker (1974), McInnes (1981), Murdin (1985), Munoz-Tunon and Fuentes (1990), or Fuentes and Munoz-Tunon (1990). While a good comparison study is hard to find, the Italian Galileo Telescope Project Phase A Report suggests that Mauna Kea is superior to La Palma. Anecdotal evidence also suggests that there appears to be a large local variation in the site properties on Canary Islands, depending on the topography, and atmospheric dust from Africa may present problems in terms of the photometry and laser backscatter.

Other known areas include Antarctica (which provides some special purpose advantages for the IR to mm wavelengths), Maidanak in Uzbekistan, Oukaimeden in Morocco, and others. Namibia seems to offer a number of potentially good sites, including Gamsberg. See the ESO Web site for links and more details. However, none of these sites seem suitable for CELT for a variety of reasons.

13.5 Site Selection Process

The preceding discussion gives an indication of the complexity of the issues associated with the site selection. The scope of the problem is magnified by the short time available to make the CELT site decision.

To summarize, we presently consider three plausible venues for CELT:

1. Mauna Kea, where the critical issue is whether the designated ELT site in the MK Master Plan is at all competitive.

2. Chile, where a number of viable candidate sites are being identified by the NOAO and ESO studies, and will presumably be tested in more detail.

3. Possible sites in the North America based on a search in the southwest U.S. and northern Mexico.

Therefore, we recommend the following:

1. Further strengthening of collaborative ties with the ELT site selection groups at NOAO and ESO, including joint projects and data exchanges for mutual benefit. This may be especially valuable for the site testing in Chile. Work on Mauna Kea requires collaboration with the Institute for Astronomy at University of Hawaii, and may also involve collaboration with the Gemini/NOAO groups. Work in Mexico would require a collaboration with their astronomical community.

2. Constructing a short list of candidate sites, comparing them using satellite-based studies and other available data, and selecting a first, second, and third choice site in Chile and North America. These should then be compared quantitatively with Mauna Kea. To this purpose, we recommend completing the satellite data studies of interesting areas in North America and on Mauna Kea, comparable to the study already done for Chile by Dr. A. Erasmus, a consulting astrometeorologist. These studies should be completed in CY 2002 Additional astroclimatological studies may be commissioned if deemed necessary.

3. Implementation of intensive site testing campaigns on Mauna Kea, in Chile, and in North America. On Mauna Kea, the key question is the comparison of the designated ELT site area on the Northern shield with the summit and ridge sites. In Chile and in North America, the top sites will be selected from the satellite studies. We propose that the Mauna Kea and the Chile campaigns are started first and conducted in parallel, with the two locations on Mauna Kea and two sites in Chile tested at the same time. North America sites would be added as the work progresses. These studies may start in late CY 2002 or early 2003, and last through early to mid-CY 2006. They will require considerable equipment purchases and hiring of the necessary personnel.

4. Perform numerical atmospheric modeling of selected sites, in collaboration with the groups at NOAO and in Mexico.

5. Reach the final site decision by mid-CY 2006. Following the formal site acquisition, initiate a long-term atmospheric characterization program on this site.

As the specific initial steps in CY 2002, we recommend the following:

1. Complete the satellite data study of North America by Dr. Andre Erasmus, through a consultant contract, in collaboration with the NOAO group. In progress, expected completion by mid-CY 2002.

2. Commission a similar study from Dr. Erasmus to provide a uniform and quantitative comparison of Mauna Kea and the sites in Chile and North America, again in collaboration with the NOAO group.

3. Develop and formalize the necessary collaborations with other groups.

4. Define and order the necessary equipment for site testing. This will include automated DIMM+MASS devices, acoustic sounders (or equivalent devices) to measure the C_n^2 profiles, meteo stations, and possibly other instruments, all of which should be designed and built to work unattended at remote locations for periods of at least of a few weeks. They should be able to survive the prevailing weather conditions year round, and have a usable lifetime expectancy of at least several years. Balloon sonde campaigns may be conducted at a few sites as well.

5. Initiate a hiring process for the necessary personnel.

6. Initiate the permit acquisition and other local arrangements so that the site testing can commence as soon as the equipment is available.

7. Start the in situ testing on Mauna Kea and the two top choice sites in Chile.

This outline is subject to modification as further work progresses on the site selection. In the rapid and aggressive schedule we envision, flexibility in response to the data and other developments is essential.

As supporting activities, we may wish to organize one or more workshops on site testing issues for ELTs, and have a visitor program whereby we bring in experts from other groups who can help us with our data analysis and planning.

All of the proposed activity should be done in collaboration and coordination with other groups, e.g., NOAO/Gemini, UH/IfA, and ESO, and include cost sharing as much as possible.

Useful Web Resources:

Site Testing and Selection:

CTIO/NOAO ELT Site Testing Program: http://www.ctio.noao.edu/sitetests/
ESO Astroclimatology Web site: http://www.eso.org/gen-fac/pubs/astclim/
ESO Search for Potential Astronomical Sites: http://www.eso.org/gen-fac/pubs/astclim/espas/
Namibia at ESO Web site: http://www.eso.org/gen-fac/pubs/astclim/espas/gamsberg/
Chajnantor at ESO Web site: http://www.eso.org/gen-fac/pubs/astclim/espas/radioseeing/
ALMA at ESO Web site: http://www.eso.org/projects/alma/
ALMA at NRAO Web site: http://www.tuc.nrao.edu/mma/sites/

Mauna Kea:

Climatology and Current Conditions: http://hokukea.soest.hawaii.edu
Master Plan: http://www.hawaii.edu/maunakea/

Night Sky Brightness:
Cinzano, et al., Web site: http://www.inquinamentoluminoso.it/dmsp/index.html
DMSP Lights on Earth: http://www.ngdc.noaa.gov/dmsp/dmsp.html

Climatology Sites:

Int'l Panel on Climate Change: http://www.ipcc.ch
NOAA: http://www.noaa.gov/
NCDC: http://www.ncdc.noaa.gov/
NCAR: http://www.ncar.ucar.edu/
EOS Visible Earth: http://visibleearth.nasa.gov/
NASA EOSDIS: http://spsosun.gsfc.nasa.gov/New_EOSDIS.html
NASA GISS: http://www.giss.nasa.gov/data/

Conference Proceedings:

"Site Testing for Future Large Telescopes." eds. A. Ardeberg and L. Woltjer, ESO CWP No. 18. 1984.
"Identification, Optimization, and Protection of Optical Telescope Sites." eds. R. Millis, et al., Lowell Observatory.1987.
"Astronomical Site Evaluation in the Visible and Radio Range", eds. J. Vernin et al., ASPCS, in press. 2002.

REFERENCES

Cinzano, P., F. Falchi, and C. Elvidge. 2001a. *Monthly Notes of the Royal Astronomical Society* in press (astro-ph/0011310).

Cinzano, P., F. Falchi, and C. Elvidge. 2001b. *Monthly Notes of the Royal Astronomical Society* in press (astro-ph/0108052).

De Young, D., and R. Charles. 1995. *Astrophysical Journal* **110**, 3107.

Echevarria, et al. 1998. *Review of Mexican Astronomy and Astrophysics* **34**, 47.

Ehgamberdiev, S., et al. 2000. *Astronomy and Astrophysics Supplement* **145**, 293.

Erasmus, A., and B. Barnes. 1989. Unpublished IfA (Univ. of Hawaii) technical report.

Erasmus, A. 2000. "A Satellite Survey of Water Vapor and Cloud Cover at Selected Existing and Potential Infrared Telescope Sites in the Southwestern U.S.A." Rocky Mountains Observatories Consortium.

Erasmus, A., and C. van Staden. 2001. "A Satellite Survey of Cloud Cover and Water Vapor in Northern Chile." Prepared for CTIO and University of Tokyo.

Fuentes, F., and C. Munoz-Tunon. 1990. *Astrophysics and Space Science* **171**, 267.

Garstang, R. 1989a. *Annual Review of Astronomy and Astrophysics* **27**, 19.

Garstang, R. 1989b. *Publication of the Astronomical Society of the Pacific* **101**, 306.

Garstang, R. 1991. *Publication of the Astronomical Society of the Pacific* **103**, 1109.

Giovanelli, R., et al. 2001a. *Publication of the Astronomical Society of the Pacific* **113**, 789.

Giovanelli, R., et al. 2001b. *Publication of the Astronomical Society of the Pacific* **113**, 803.

Lynds, R., and J. Goad. 1984. *Publication of the Astronomical Society of the Pacific* **96**, 750.

McInnes, B., and M. Walker. 1974. *Publication of the Astronomical Society of the Pacific* **86**, 529.

McInnes, B. 1981. *Quarterly Journal of the Royal Astronomical Society* **22**, 266.

Merrill, K., G. Favot, D. Forbes, D. Morse, and G. Poczulp. 1986. SPIE Proceedings **628**, 125.

Munoz-Tunon, C., and F. Fuentes. 1990. *Astrophysics and Space Science* **171**, 257.

Murdin, P. 1985. *Vistas in Astronomy* **28**, 449.

Sarazin, M., and F. Roddier. 1990. *Astronomy and Astrophysics* **227**, 294.

Sarazin, M. 2002. EAO Search for Potential Astronomical Sites Report 1.1. Available at http://www.eso.org/gen-fac/pubs/astclim/espas/espas_reports/

Teare, S., L. Thompson, M.C. Gino, and K Palmer. 2000. *Publication of the Astronomical Society of the Pacific* **112**, 1496.

Tokovinin, A. 1998. *Astronomy Letter* **24**, 662.

Walker, M. 1970. *Publication of the Astronomical Society of the Pacific* **82**, 672.

Walker, M. 1971. *Publication of the Astronomical Society of the Pacific* **83**, 401.

Walker, M. 1983. *Publication of the Astronomical Society of the Pacific* **95**, 903.

Ziad, A., R. Conan, A. Tokovinin, F. Martin, and J. Borgnino. 2000. *Applied Optics* **39**, 5415.

Chapter 14. Computer Software and Hardware

14.1 Introduction

Computer hardware and software were not formally addressed during the CELT conceptual design phase. We present here a preliminary discussion of the major aspects of software development for the observatory. A formal development plan will be created during the next design phase. Staffing and budget estimates for this phase are given in CELT Report No. 35 (Nelson and Mast, eds., 2002).

At this early stage it is difficult to quantify the overall CELT project cost attributable to software. For projects of this complexity the cost of software inefficiencies and rework can be a significant fraction of the total effort. One such study from the Standish Group (Royce 1998) showed only a small fraction of projects were delivered successfully on schedule and within initial budget estimates. Effective software management will determine success or failure.

Software development methodologies vary widely. At one end of the spectrum is the highly formalized (and expensive) full CMM Level Five compliance (Software Engineering Institute Capability Maturity Model) featuring a highly structured division of labor in a large software team. Near the other end is an informal approach, in which the same engineers who develop software also participate in the integration, testing, and on-going operations of their subsystems. The latter approach was used for software development at the Keck Observatory with mixed results. We recommend for CELT a more formal approach than that used for Keck. This would include the dedication of resources to requirements analysis, architecture, and design phases prior to software implementation and testing. We believe that this approach will result in substantial savings in maintenance and rework costs.

There is no simple solution for software management, which to a large extent is only learned from hands-on, real-world experience. Nonetheless, we outline here some steps towards a coherent software implementation for CELT. This is only the framework for a modern software management process to be defined for CELT and fully elaborated in the next design phase of the project.

14.2 Lessons Learned at the Keck Observatory

We summarize here some of the major lessons learned at Keck. Based thereon, we make recommendations for reducing the risks and costs of the CELT software.

1. Development infrastructure
    a. Invest in tools to analyze software metrics and their trends.
    b. Invest in automated problem reporting and tracking.
    c. Invest in computer and network resource monitoring and alarms.
    d. Invest in automated calibration tools (i.e., pointing test software, encoder calibration, telescope balance, etc.). Ensure these are delivered early in the development cycle, well before hardware integration and test.

2. Design for testing, integration, and maintenance
    a. Make test planning a part of the initial design.
    b. Utilize independent software testers.
    c. Provide subsystem simulators to avoid tying up telescope and instruments.

3. Apply principles of interactive design (Cooper 1995).

4. Utilize design patterns (Gamma, et. al., 1995).

5. Ensure that the infrastructure meets the needs of, and is utilized by, application developers.

6. Modularize GUI development apart from underlying observatory control functionality. This allows GUI development to evolve independently, to take advantage of rapidly changing technologies.

7. Control all subsystems (instrument and observatory) from an external application, using the common application programming interfaces.

14.3 Software Functional Requirements

Detailed functional requirements for the observatory computing software and hardware resources will be developed during the preliminary design phase. We list the major categories of functional requirements.

Observatory Control

1. Telescope control system
    - 1.1 Acquisition
    - 1.2 Pointing and tracking
    - 1.3 Autoguiding and wavefront sensing
    - 1.4 Primary mirror active figure control
    - 1.5 Secondary mirror positioning and active figure control
    - 1.6 Tertiary mirror pointing
    - 1.7 Optics alignment and calibration
    - 1.8 Safety interlock monitoring and control

   (Each telescope control subsystem will require software for configuration, control, and real-time diagnostics.)

2. Enclosure, shutter, and windscreen motion control

3. Environment monitoring and control
    - 3.1 Site meteorology (e.g. cloud cover, humidity)
    - 3.2 Seeing condition monitoring (e.g., $C_n^2(h)$, vertical wind profiles)
    - 3.3 Dome and telescope temperature monitoring
    - 3.4 Dome and telescope ventilation/temperature control
    - 3.5 Potential need for sodium layer column density and/or aerosol backscatter monitoring

4. Adaptive Optics (AO)
    - 4.1 AO system configuration
    - 4.2 AO system control
    - 4.3 Real-time wavefront processing
    - 4.4 AO procedure automation (e.g. calibration, registration)
    - 4.5 Laser control and safety automation (e.g. human, aircraft safety)
    - 4.6 Real-time telemetry analysis and performance diagnostics
    - 4.7 Non-common-path wavefront metrology

5. Instruments
   - 5.1 Instrument configuration
   - 5.2 Instrument control
   - 5.3 Instrument procedure automation (e.g., dither patterns)
   - 5.4 Real-time diagnostics (e.g., cryogen status)
   - 5.5 Detector control and readout (We envision standardized detector readout controls, based upon the current ASTEROID controller development effort.)

6. Operations execution
   - 6.1 GUI and scripted operations
   - 6.2 Remote observing
   - 6.3 Queue scheduling and autonomous operations
   - 6.4 Instrument detector quick-look display

Data Management

7. Pre-Observation
   - 7.1 Proposal preparation tools
   - 7.2 AO observation planning tools
   - 7.3 Observing sequence planning tools
   - 7.4 Weather forecasting

8. Post-Observation
   - 8.1 Data quality monitoring and archiving
     - 8.1.1 Storage, cataloguing, retrieval, distribution
     - 8.1.2 Engineering support and safety backups
     - 8.1.3 Long-term science support
   - 8.2 Data reduction pipelines
     - 8.2.1 AO point spread function (PSF) estimation
     - 8.2.2 High-contrast PSF calibration
     - 8.2.3 Instrument-specific reduction pipelines

Software Development Tools and Infrastructure

9. Subsystem simulators
   - 9.1 Telescope and enclosure
   - 9.2 Primary mirror
   - 9.3 AO system and instrument interfaces
   - 9.4 Artificial light source
   - 9.5 Motor control
   - 9.6 Other subsystem simulators defined with clean interfaces via the architecture development, as the design process evolves

10. Computing infrastructure
   - 10.1 Documentation of tools and standards
   - 10.2 Intranet configuration management
   - 10.3 Telemetry logging and alarms
   - 10.4 Real-time control system framework (EPICS is used at Keck, for example)
   - 10.5 Common application programming interface (API) to subsystems (KTL is used at Keck)

10.6 Databases (e.g., day and night logs, equipment configurations, schedules, engineering change requests, inventory, drawings)

11. Communications
    - 11.1 Inter-process communications
    - 11.2 Inter-subsystem communications (e.g., between the telescope, AO, and instrument subsystems)
    - 11.3 Inter-site communications (e.g., between summit, headquarters, remote observing stations)

    We must address both computer and human communications requirements.

Administrative

12. Scheduling and support
    - 12.1 Observatory personnel scheduling
    - 12.2 Maintenance scheduling
    - 12.3 Business support (e.g., payroll, contracting, procurement)
    - 12.4 Desktop computing support (e.g., headquarters staff)
    - 12.5 Public relations and science dissemination

14.4 Software Development Process

Although a description of an integrated software management plan for CELT is outside the scope of this report, we can give a set of guidelines for the next phases. These software engineering principles define a management process that might be used throughout the CELT project life. We also discuss, within this framework, important software issues that must be addressed during the next design phase before proceeding to full-scale development.

14.4.1 Modern Software Engineering

Many software projects still use the conventional *waterfall model* in which the development process evolves in a mostly sequential progression from requirements analysis, through design, implementation, and deployment. Using this model, the investment (and cost of rework) steadily increases through the lifetime of the project. This model and its variations are no longer considered to be an appropriate framework to manage the complexity of large-scale software systems. To avoid the risks and shortcomings of these traditional approaches, we recommend an alternative *iterative* software management process based on the following proven principles,

1. Use an iterative and incremental software development process.
   Rather than progressing sequentially, development proceeds as a series of iterations, starting with an initial emphasis on refining the core requirements and defining a core software architecture. The process then continues building on the core architecture with incremental build releases until all the required levels of functionality and performance are achieved.

2. Focus development on an architecture-driven approach.
   The software architecture is the central design element of a system. An early focus on constructing executable subsets or prototypes of the architecture, with the goal of converging on a stable baseline, allows design risks to be recognized and resolved early before committing to full-scale development.

3. Use *concurrent engineering* to accelerate development cycles.
   Concurrent engineering complements the iterative development process by allowing activities in each phase to be performed concurrently, in varying proportions and with varying level of efforts. For example, architectural design can be performed in parallel with the specification of requirements as part of the engineering phase of the project. Similarly, test planning and preparation can be performed in parallel with design and development activities.

4. Use *demonstration-based reviews* (or milestones) to get early user's evaluations and performance feedback.
   The traditional emphasis on document-based reviews and inspections is shifted to a focus on assessment through demonstrations. Demonstrations of increasing system capabilities should focus on critical elements of the operational scenarios as the architecture evolves.

5. Maximize *component-based development* where it is architecturally feasible in order to reduce complexity and minimize overall development costs.
   Component is used here in the broadest sense to include both commercial components, such as operating systems, database management systems, windowing environments, and networking; and custom-developed components, such as common architectures. While this principle may sound obvious, it has not been widely embraced and integrated into the development process in practice.

These five principles together form the essence of an iterative development process. An early implementation focus on the software architecture, through iterative refinement of core requirements, allows critical design risks to be resolved early. The process also enables activities in each phase, and across various teams, including observatory, instrument groups, etc., to evolve concurrently, all tied together through a common architecture baseline.

14.4.2 Software and Computing Hardware Requirements Analysis

The specification of requirements is a difficult and important part of the software development process for CELT. The iterative process proposed here does not require complete and unambiguous specification of the requirements before allowing other phases to begin. Instead, the project initially will focus on and define the core requirements, refining them in one or more iterations with the goal of enabling early construction of an architecture prototype. The complete definition of the requirements will be achieved in later phases.

Another important objective in the requirements phase is to refine the operations concept for CELT presented in CELT Report No. 35 (Nelson and Mast, eds., 2002). Specifying key operations scenarios will be required to make trade-offs in defining the software architecture (as well as the hardware architecture). Decisions about CELT operations requirements affect the way in which the architectural design proceeds and vice-versa. It is crucial that systems engineering activity be closely coordinated with the software architecture team.

As one example, support for remote observing has typically been an afterthought at existing observatories, resulting in remote-observing facilities with limited capabilities and marginal performance. In order to efficiently support remote observing, CELT's software architecture will need to be fully distributed and client-server based, with instrument-control applications clearly separated from their graphical user interface.

Changes to the software architecture to support late operations requirements can be disruptive and prohibitively costly, if not impossible. The *software architect* is responsible for coordinating the requirements and design trade-offs, and must work closely with the telescope, AO, and instrument systems engineers.

During this phase, it is also necessary to establish an integrated development environment (IDE) sufficient to support a level of automation of the iterative development process. This includes computer-aided software engineering (CASE) tools to support requirements analysis, design modeling, host/target cross-development, automated regression testing, document generation, etc. Visual modeling tools, such as the Unified Modeling Language (UML), will be evaluated and tailored for use within CELT to provide a standard and rigorous means of communicating requirements and design among the different development teams.

14.4.3 Software Architecture and Design

An architecture-driven iterative development process provides an early focus on the most critical design aspect of the CELT software, namely its architecture. This architecture consists of

- infrastructure, control elements, communications, and data interfaces that allow all subsystems (e.g., telescope, AO, instruments, etc.) to cooperate as an integral system;
- a common framework for the various software development groups to cooperate effectively as a cohesive team. The design and early construction of executable iterations of the evolving architecture allow it to be validated throughout the engineering phase.

Major design and cost risks are addressed early to avoid downstream surprises and potentially expensive rework in later phases. Establishing an early software architecture baseline also forces agreement on all the subsystem's external interfaces and their critical internal interfaces.

There are many sources of architectural risk for CELT, which must be mitigated during this phase through the development of disposable prototypes or exploratory testbeds. Some of the more important sources include system performance, system reliability, and adaptability to change. Sources of engineering risk include the many challenges associated with a 30-meter primary mirror, instrument control, numerous unproven technologies, and advanced computing platforms required by CELT AO.

The choice of commercial components over custom components will also have a significant impact on CELT's overall architecture. Selecting the right technology to integrate into the software architecture will be difficult in a rapidly evolving industry, and is widely known as a potential make-or-break decision for a project. The selection will require intimate knowledge of the prevalent technologies, and an understanding and analysis of the relevant requirements for the system being built. Given the long development time and long-lived nature of CELT, the risk of technology obsolescence for the delivered software system throughout its life span must be carefully assessed against its adaptability to change. Of equal importance, the architecture must be developed to minimize development and cost, including the impact of the software on subsequent maintenance and operations costs of the observatory.

14.4.4 Reducing Operations Cost

One of the principal objectives of the CELT software is to reduce both recurring day-to-day operations costs and non-recurring implementation costs for the observatory by changing the way the telescope and instruments are configured and controlled, and improving the way operations capabilities are

implemented. The latter can be achieved by using a software management process like the one described here and leveraging architectural commonality across subsystems where feasible. The former is only possible by identifying and directly supporting in the software architecture all the required aspects of observatory operations. Observatory operations are described in Volume Two, Chapter 7, which includes real-time operations, remote observing, operations support, maintenance, and sustaining activities.

14.4.5 Software Implementation and Test

In contrast to the conventional practice, software implementation and test activities are to be initiated early in CELT's design and development. While the iterative development process presented here promotes early and continuous demonstrations of executable subsets of the architecture, the project will only commit to full-scale implementation and testing of all the subsystems *after* the requirements and management plans are stabilized and the architecture baselined.

The continuous software integration and testing inherent in the iterative development approach not only enable early detection of design flaws, they also provide more timely insight into important design trade-off and performance issues. This approach avoids the potentially disruptive *big-bang* integration that can lead to costly, late risk resolution. As mentioned above, an independent test team will perform integration and testing to minimize ownership bias and maximize concurrency of activities.

14.4.6 Recommendations for Personnel and Organization

Projects of CELT's scale require a specific human organization to successfully undertake software development. In the development model employed at Keck Observatory, the software for each instrument and the AO system was independently designed, budgeted, and managed by the group building the associated hardware. This has led to difficult and costly integration activities, wherein the observatory must support the instrument principal investigator choice of platform, language, etc.

For CELT, we recommend an organization that centralizes the high-level software architectural effort and establishes functional interfaces iteratively throughout the project. This requires a significant software team for the initial stages of the project design phase, consisting of the following key positions:

1. A *software project manager*, reporting directly to the CELT construction project manager, to provide overall project control and management responsibilities, initial budget and schedule estimates, and to fully elaborate a software management plan and work breakdown structure based on the software engineering process recommended here.

2. A *systems engineer*, to coordinate systems engineering activities between the various groups (observatory, AO, and instruments), and be responsible for software requirements specifications and observatory operations concept.

3. A *software architect*, to lead the overall architectural design activity, including all demonstration planning and coordination, and be responsible for all requirements and design trade-offs.

4. A *software engineer*, to provide the lead support to the software architecture team for all prototyping activities.

We also recommend that a software working group be formed at the start of CELT's next design phase to address, among other things, system-wide operability and software issues, and to define the critical elements of the operations. This working group will also provide a forum for technical interchange with other similar telescope facilities. Preliminary cost and schedule estimates for software development during the CELT design phase are presented in CELT Report No. 35 (Nelson and Mast, eds.).

REFERENCES

Boehm, Barry, et al. 2000. *Software Cost Estimation with COCOMO II*. Prentice Hall, ISBN 0-13-026692-2.

Cooper, Alan. 1995. *About Face: The Essentials of User Interface Design*. IDG Books.

Fowler, Martin and Kendall Scott. 1999. *UML Distilled*. Second Edition, Addison Wesley, ISBN 0-201-65783-X, Chapter 2.

Gamma, E., R. Helm, R. Johnson, and J. Vlissides. 1995. *Design Patterns: Elements of Reusable Object-Oriented Software*. Reading, MA: Addison-Wesley.

Nelson, J. and T. Mast, eds. 2002. "Operations." CELT Report No. 35.

Royce, Walker. 1998. *Software Project Management: A Unified Framework*. Addison-Wesley.